Okbay, A., Baselmans, B. M. L., De Neve, J-E., Turley, P., Nivard, M. G., Fontana, M. A., Meddens, S. F. W., Linnér, R. K., Rietveld, C. A., Derringer, J., Gratten, J., Lee, J. J., Liu, J. Z., de Vlaming, R., Ahluwalia, T. S., Buchwald, J., Cavadino, A., Frazier-Wood, A. C., Furlotte, N. A. (2016). Genetic variants associated with subjective well-being, depressive symptoms, and neuroticism identified through genome-wide analyses. *Nature Genetics*, *48*(6), 624-633. https://doi.org/10.1038/ng.3552

Peer reviewed version

Link to published version (if available):
10.1038/ng.3552

Genetic Associations with Subjective Well-Being Also Implicate Depression and Neuroticism

Authors: All authors and their affiliations appear at the end of the paper*

*Correspondence to: Daniel Benjamin (djbenjam@usc.edu), Meike Bartels (m.bartels@vu.nl), or Philipp Koellinger (p.d.koellinger@vu.nl)

ABSTRACT: We conducted genome-wide association studies of three phenotypes: subjective well-being (SWB; N = 298,420), depressive symptoms (DS; N = 161,460), and neuroticism (N = 170,910). We identified three variants associated with SWB, two with DS, and eleven with neuroticism, including two inversion polymorphisms. The two DS loci replicate in an independent depression sample. Joint analyses that exploit the high genetic correlations between the phenotypes ($|\hat{\rho}| \approx 0.8$) strengthen the overall credibility of the findings, and allow us to identify additional variants. Across our phenotypes, loci regulating expression in central nervous system and adrenal/pancreas tissues are strongly enriched for association.

Subjective well-being (SWB)—as measured by survey questions on life satisfaction, positive affect, or happiness—is a major topic of research within psychology, economics, and epidemiology. Twin studies have found that SWB is genetically correlated with depression (characterized by negative affect, anxiety, low energy, bodily aches and pains, pessimism, and other symptoms) and neuroticism (a personality trait characterized by easily experiencing negative emotions such as anxiety and fear)[1–3]. Depression and neuroticism have received much more attention than SWB in genetic-association studies, but the discovery of associated genetic variants with either of them has proven elusive[4,5].

In this paper, we report a series of separate and joint analyses of SWB, depressive symptoms (DS), and neuroticism. Our primary analysis is a genome-wide association study (GWAS) of SWB based on data from 59 cohorts (N = 298,420). This GWAS identifies three loci associated with SWB at genome-wide significance ($p < 5\times10^{-8}$). We supplement this primary analysis with auxiliary GWAS meta-analyses of DS (N = 180,866) and neuroticism (N = 170,910), performed by combining publicly available summary statistics from published studies with new genome-wide analyses of additional data. In these auxiliary analyses we identify two loci associated with DS and eleven with neuroticism, including two inversion polymorphisms. In depression data from an independent sample (N = 368,890), both DS associations replicate (p = 0.004 and p = 0.015).

In our two joint analyses, we exploit the high genetic correlation between SWB, DS, and neuroticism (i) to evaluate the credibility of the 16 genome-wide significant associations across the three phenotypes, and (ii) to identify novel associations (beyond those identified by the GWAS). For (i), we investigate whether our three SWB-associated SNPs "quasi-replicate" by testing them for association with DS and neuroticism. We similarly examine the quasi-replication record of the DS and neuroticism loci by testing them for association with SWB. We find that the quasi-replication record closely matches what would be expected given our statistical power if none of the genome-wide significant associations were chance findings. These results strengthen the credibility of (most of) the original associations. For (ii), we use a "proxy phenotype" approach[6]: we treat the set of loci associated with SWB at $p < 10^{-4}$ as candidates, and we test them for association with DS and neuroticism. At the Bonferroni-adjusted 0.05 significance threshold, we identify two loci associated with both DS and neuroticism and another two associated with neuroticism.

In designing our study, we faced a tradeoff between analyzing a smaller sample with a homogeneous phenotype measure versus attaining a larger sample by jointly analyzing data from multiple cohorts with

heterogeneous measures. For example, in our analysis of SWB, we include measures of both life satisfaction (LS) and positive affect (PA), even though these constructs are conceptually distinct[7]. In **Supplementary Note**, we present a theoretical framework for evaluating the costs and benefits of pooling heterogeneous measures. In our context, given the high genetic correlation across measures, the framework predicts that pooling increases statistical power to detect variants. This prediction is supported by our results.

RESULTS

GWAS of SWB

Following a pre-specified analysis plan, we conducted a sample-size-weighted meta-analysis (N = 298,420) of cohort-level GWAS summary statistics. The phenotype measure was LS, PA, or (in some cohorts) a measure combining LS and PA. We confirmed previous findings[9] of high pairwise genetic correlation between LS and PA using bivariate LD Score regression[10] ($\hat{\rho}$ = 0.981 (SE = 0.065); **Supplementary Table 1**). Details on the 59 participating cohorts, their phenotype measures, genotyping, quality-control filters, and association models are provided in Online Methods, **Supplementary Note**, and **Supplementary Tables 2-6**.

As expected under polygenicity[11], we observe inflation of the median test statistic (λ_{GC} = 1.206). The estimated intercept from LD Score regression (1.012) suggests that nearly all of the inflation is due to polygenic signal rather than bias. We also performed family-based analyses that similarly suggest minimal confounding due to population stratification (Online Methods). Using a clumping procedure (Online Methods), we identified three approximately independent SNPs reaching genome-wide significance ("lead SNPs"). These three lead SNPs are indicated in the Manhattan plot (**Figure 1a**) and listed in **Table 1**. The SNPs have estimated effects in the range 0.015 to 0.018 standard deviations (SDs) per allele (each $R^2 \approx$ 0.01%).

We also conducted separate meta-analyses of the components of our SWB measure, LS (N = 166,205) and PA (N = 180,281) (Online Methods). Consistent with our theoretical conclusion that pooling heterogeneous measures increased power in our context, the LS and PA analyses yielded fewer signals across a range of p-value thresholds than our meta-analysis of SWB (**Supplementary Table 7**).

GWAS of DS and neuroticism

We conducted auxiliary GWAS of DS and neuroticism (see Online Methods, **Supplementary Note**, and **Supplementary Tables 8-12** for details on cohorts, phenotype measures, genotyping, association models, and quality-control filters). For DS (N = 180,866), we meta-analyzed publicly available results from a study performed by the Psychiatric Genomics Consortium (PGC)[12] together with new results from analyses of the initial release of the UK Biobank data (UKB)[13] and the Resource for Genetic Epidemiology Research on Aging (GERA) Cohort[14]. In UKB (N = 105,739), we constructed a continuous phenotype measure by combining responses to two questions, which ask about the frequency with which the respondent experienced feelings of unenthusiasm/disinterest or unenthusiasm/disinterest in the past two weeks. The other cohorts had ascertained case-control data on major depressive disorder (GERA: N_{cases} = 7,231, $N_{controls}$ = 49,316; PGC: N_{cases} = 9,240, $N_{controls}$ = 9,519).

For neuroticism (N = 170,910), we pooled summary statistics from a published study by the Genetics of Personality Consortium (GPC)[4] with results from a new analysis of UKB data. The GPC (N = 63,661) harmonized different neuroticism batteries. In UKB (N = 107,245), our measure was the respondent's score on a 12-item version of the Eysenck Personality Inventory Neuroticism scale[15].

In both the DS and neuroticism GWAS, the heterogeneous phenotypic measures are highly genetically correlated (**Supplementary Table 1**). As in our SWB analyses, there is substantial inflation of the median

test statistics (λ_{GC} = 1.168 for DS, λ_{GC} = 1.317 for neuroticism), but the estimated LD Score intercepts (1.008 and 0.998, respectively) suggest that bias accounts for little or none of the inflation.

For DS, we identified two lead SNPs, indicated in the Manhattan plot (**Fig. 1b**). For neuroticism, our meta-analysis yielded 16 loci that are independent according to our locus definition (**Fig. 1c**). However, 6 of these reside within a well-known inversion polymorphism on chromosome 8[16]. We established that all genome-wide significant signals in the inversion region are attributable to the inversion, and we confirmed that the inversion is associated with neuroticism in both of our neuroticism datasets, the GPC and the UKB (Online Methods and **Supplementary Note**). In our list of lead SNPs (**Table 1**), we only retain the most strongly associated SNP from these 6 loci to tag the chromosome 8 inversion.

Another lead SNP associated with neuroticism, rs193236081, is located within a well-known inversion polymorphism on chromosome 17. We established that this association is attributable to the inversion polymorphism (Online Methods and **Supplementary Note**). Because this inversion yields only one significant locus and is genetically complex[17], we hereafter simply use its lead SNP as its proxy. Our neuroticism GWAS therefore identified 11 lead SNPs, two of which tag inversion polymorphisms. A concurrent, unpublished neuroticism GWAS using a subset of our sample reports similar findings[18].

As shown in **Table 1**, the estimated effects of all lead SNPs associated with DS and neuroticism are in the range 0.020 to 0.031 SDs per allele ($R^2 \approx 0.02\%$ to 0.04%). In the UKB cohort we estimated the effect of an additional allele of the chromosome 8 inversion polymorphism itself on neuroticism to be 0.035 SDs (**Supplementary Table 13**).

Genetic overlap across SWB, DS, and neuroticism

Figure 2a shows that the three pairwise genetic correlations between our phenotypes, estimated using bivariate LD Score regression[10], are substantial: -0.81 (SE = 0.046) between SWB and DS, -0.75 (SE = 0.034) between SWB and neuroticism, and 0.75 (SE = 0.027) between DS and neuroticism. Using height as a negative control, we also examined pairwise genetic correlations between each of our phenotypes and height and, as expected, found all three to be modest, e.g., 0.07 with SWB (**Supplementary Table 1**). The high genetic correlations between SWB, DS, and neuroticism may suggest that the genetic influences on these phenotypes are predominantly related to processes common across the phenotypes, such as mood, rather than being phenotype-specific.

Quasi-replication and Bayesian credibility analyses

We assessed the credibility of our findings using a standard Bayesian framework[19,20] in which a positive fraction of SNPs have null effects and a positive fraction have non-null effects (Online Methods). For each phenotype, the non-null effect sizes are assumed to be drawn from a normal distribution whose variance is estimated from the GWAS summary statistics. As a first analysis, for each lead SNP's association with its phenotype, we calculated the posterior probability of null association after having observed the GWAS results. We found that, for any assumption about the fraction of non-null SNPs in the range 1% to 99%, the probability of true association always exceeds 95% for all 16 loci (and always exceeds 98% for 14 of them).

To further probe the credibility of the findings, we performed "quasi-replication" exercises (Online Methods) in which we tested the SWB lead-SNPs for association with DS and neuroticism. We similarly tested the DS lead-SNPs and the neuroticism lead-SNPs for association with SWB. Below, we refer to the phenotype for which the lead SNP was identified as the first-stage phenotype and the phenotype used for the quasi-replication as the second-stage phenotype. To avoid sample overlap, for each quasi-replication analysis we omitted any cohorts that contributed to the GWAS of the first-stage phenotype.

Results of the quasi-replication of the three SWB lead-SNPs are shown in **Figure 3a**. For ease of interpretation, the reference allele for each association in the figure is chosen such that the predicted sign of the second-stage estimate is positive. We find that two out of the three SWB lead-SNPs are significantly associated with DS ($p = 0.004$ and $p = 0.001$) in the predicted direction. For neuroticism, where the second-stage sample size ($N = 68{,}201$) is about half as large, the SWB-increasing allele has the predicted sign for all three SNPs, but none reach significance.

Figures 3b and 3c show the results for the DS and neuroticism lead-SNPs, respectively. In each panel, the blue crosses depict results from the quasi-replications where SWB is the second-stage phenotype. We find that the two DS lead-SNPs have the predicted sign for SWB, and one is nominally significant ($p = 0.04$). Finally, of the eleven neuroticism lead-SNPs, nine have the predicted sign for SWB. Four of the eleven are nominally significantly associated with SWB, all with the predicted sign. One of the four is the SNP tagging the inversion on chromosome 8[16]. That SNP's association with neuroticism (and likely with SWB) is driven by its correlation with the inversion (**Supplementary Fig. 1**).

To evaluate what these quasi-replication results imply about the credibility of the 16 GWAS associations, we compared the observed quasi-replication record to the quasi-replication record expected given our statistical power. We calculated statistical power using our Bayesian framework, under the hypothesis that each lead SNP has a non-null effect on both the first- and second-stage phenotypes. Our calculations take into account both the imperfect genetic correlation between the first- and second-stage phenotypes and inflation of the first-stage estimates due to the well-known problem of winner's curse (Online Methods). Of the 19 quasi-replication tests, our calculations imply that 16.7 would be expected to yield the anticipated sign and 6.9 would be significant at the 5% level. The observed numbers are 16 and 7. Our quasi-replication results are thus consistent with the hypothesis that none of the 16 genome-wide significant associations are chance findings, and in fact strengthen the credibility of our GWAS results (**Supplementary Table 14**).

Lookup of DS and neuroticism lead-SNPs

Investigators of an ongoing large-scale GWAS of major depressive disorder ($N = 368{,}890$) in the 23andMe cohort shared association results for the loci identified in our DS and neuroticism analyses (Online Methods and **Supplementary Table 15**)[21]. Because the depression sample overlaps with our SWB sample, we did not request a lookup of the SWB-associated SNPs.

In **Figures 3b** and **3c**, the results are depicted as green crosses. For interpretational ease, we chose the reference allele so that positive coefficients imply that the estimated effect is in the predicted direction. All 13 associations have the predicted sign. Of the 11 neuroticism polymorphisms, four are significantly associated with depression at the 5% level. Both of the DS lead-SNPs replicate ($p = 0.004$ and $p = 0.015$), with effect sizes (0.007 and -0.006 SDs per allele) strikingly close to those predicted by our Bayesian framework (0.008 and -0.006) (**Supplementary Table 14** and **Supplementary Table 15**).

Panel A of **Table 1** summarizes the results for the 16 lead SNPs identified across our separate GWA analyses of the three phenotypes. The right-most column summarizes the statistical significance of the quasi-replication and depression lookup analyses of each SNP.

Proxy-phenotype analyses

To identify additional SNPs associated with DS, we conducted a two-stage "proxy phenotype" analysis (Online Methods). In the first stage, we ran a new GWAS of SWB to identify a set of candidate SNPs. Specifically, from each locus exhibiting suggestive evidence of association ($p < 10^{-4}$) with SWB, we retained the SNP with the lowest p-value as a candidate. In the second stage, we tested these candidates for association with DS at the 5% significance threshold, Bonferroni-adjusted for the number of candidates. We used an analogous two-stage procedure to identify additional SNPs associated with neuroticism. The first-

stage SWB sample differs across the two proxy-phenotype analyses (and from the primary SWB GWAS sample) because we assigned cohorts across the first and second stages so as to maximize statistical power for the overall procedure.

For DS, there are 163 candidate SNPs. 115 of them (71%) have the predicted direction of effect on DS, 20 are significantly associated at the 5% significance level (19 in the predicted direction), and two remain significant after Bonferroni adjustment. For neuroticism, there are 170 candidate SNPs. 129 of them (76%) have the predicted direction of effect, all 28 SNPs significant at the 5% level have the predicted sign, and four of these remain significant after Bonferroni adjustment (**Supplementary Fig. 2** and **Supplementary Tables 16** and **17**). Two of the four are the SNPs identified in the proxy-phenotype analysis for DS.

Table 1 lists the four SNPs in total identified by the proxy-phenotype analyses.

Biological analyses

To shed some light on possible biological mechanisms underlying our findings, we conducted several analyses.

We began by using bivariate LD Score regression[10] to quantify the amount of genetic overlap between each of our three phenotypes and ten neuropsychiatric and physical health. **Figures 2b** and **c** display the estimates for SWB and the *negative* of the estimates for DS and neuroticism (since SWB is negatively genetically correlated with DS and neuroticism). SWB, DS, and neuroticism have strikingly similar patterns of pairwise genetic correlation with the other phenotypes.

Figure 2b shows the results for the five neuropsychiatric phenotypes we examined: Alzheimer's disease, anxiety disorders, autism spectrum disorder, bipolar disorder, and schizophrenia. For four of these phenotypes, genetic correlations with depression (but not neuroticism or SWB) were reported in Bulik-Sullivan et al.[10]. For schizophrenia and bipolar disorder, our estimated correlations with DS, 0.33 and 0.26, are substantially lower than Bulik-Sullivan et al.'s point estimates but contained within their 95% confidence intervals. By far the largest genetic correlations we estimate are with anxiety disorders: −0.73 with SWB, 0.88 with DS, and 0.86 with neuroticism. Genetic correlations estimated from GWAS data have not been previously reported for anxiety disorders.

Figure 2c shows the results for five physical health phenotypes that are known or believed to be risk factors for various adverse health outcomes: body mass index (BMI), ever-smoker status, coronary artery disease, fasting glucose, and triglycerides. The estimated genetic correlations are all small in magnitude, consistent with earlier work, although the greater precision of our estimates allows us to reject null effects in most cases. The signs are generally consistent with those of the phenotypic correlations reported in earlier work between our phenotypes and outcomes such as obesity[22], smoking[23,24], and cardiovascular health[25].

Next, to investigate whether our GWAS results are enriched in particular functional categories, we applied stratified LD Score regression[26] to our meta-analysis results. In our first analysis, we report estimates for all 53 functional categories included in the "baseline model"; the results for SWB, DS, and neuroticism are broadly similar (**Supplementary Tables 18-20**) and are in line with what has been found for other phenotypes[26]. In our second analysis, the categories are groupings of SNPs likely to regulate gene expression in cells of a specific tissue. The estimates for SWB, DS, and neuroticism are shown in **Figure 4a**, alongside height, which is again included as a benchmark[27] (see also **Supplementary Table 21**).

We found significant enrichment of CENTRAL NERVOUS SYSTEM for all three phenotypes and, perhaps more surprisingly, enrichment of ADRENAL/PANCREAS for SWB and DS. The cause of the ADRENAL/PANCREAS enrichment is unclear, but we note that the adrenal glands produce several hormones, including cortisol, epinephrine, and norepinephrine, known to play important roles in the bodily regulation of mood and stress.

It has been robustly found that blood serum levels of cortisol in patients afflicted by depression are elevated relative to controls[28].

While the above analyses utilize the genome-wide data, we also conducted three analyses (Online Methods) restricted to the 16 GWAS and four proxy-phenotype SNPs in **Table 1**. In brief, we ascertained whether each SNP (or a variant in strong linkage disequilibrium (LD) with it) falls into any of the following three classes: (i) resides in a locus for which genome-wide significant associations with other phenotypes have been reported (**Supplementary Table 22**), (ii) is nonsynonymous (**Supplementary Table 23**), and (iii) is an eQTL in blood or in one of 14 other tissues (although the non-blood analyses are based on smaller samples) (**Supplementary Table 24**). Here we highlight a few particularly interesting results.

We found that five of the 20 SNPs are in loci in which genome-wide significant associations have previously been reported. Two of these five are schizophrenia loci. Interestingly, one of them harbors the gene *DRD2*, which encodes the D_2 subtype of the dopamine receptor, a target for antipsychotic drugs[29] that is also known to play a key role in neural reward pathways[30]. Motivated by these findings, as well as by the modest genetic correlations with schizophrenia reported in **Figure 2b**, we examined whether the SNPs identified in a recent study of schizophrenia[31] are enriched for association with neuroticism in our non-overlapping UKB sample (N = 107,245). We conducted several tests and found strong evidence of such enrichment (Online Methods). For example, we found that the p-values of the schizophrenia SNPs tend to be much lower than the p-values of a randomly selected set of SNPs matched on allele frequency ($p = 6.50\times10^{-71}$).

Perhaps the most notable pattern that emerges from our biological analyses is that the inversions on chromosomes 8 and 17 are implicated consistently across all analyses. The inversion-tagging SNP on chromosome 8 is in LD with SNPs that have previously been found to be associated with BMI[32] and triglycerides[33] (**Supplementary Table 22**). We also conducted eQTL analyses in blood for the inversion itself and found that it is a significant *cis*-eQTL for 7 genes (**Supplementary Table 24**). As shown in **Figure 4b**, all 7 genes are positioned in close proximity to the inversion breakpoints, suggesting that the molecular mechanism underlying the inversion's effect on neuroticism could involve the relocation of regulatory sequences. Two of the genes (*MSRA*, *MTMR9*) are known to be highly expressed in tissues and cell types that belong to the nervous system, and two (*BLK*, *MFHAS1*) in the immune system. In the tissue-specific analyses, we found that the SNP tagging the inversion is a significant eQTL for two genes, *AF131215.9* (in tibial nerve and thyroid tissue analyses) and *NEIL2* (tibial nerve tissue), both of which are also located near the inversion breakpoint.

The SNP tagging the chromosome 17 inversion is a significant *cis*-eQTL for five genes in blood and is an eQTL in all 14 other tissues (**Supplementary Table 24**). It alone accounts for 151 out of the 169 significant associations identified in the 14 tissue-specific analyses. Additionally, the SNP is in near-perfect LD ($R^2 > 0.97$) with 11 missense variants (**Supplementary Table 23**) in three different genes, one of which is *MAPT*. *MAPT*, which is also implicated in both the blood and the other tissue-specific analyses, encodes a protein important in the stabilization of microtubules in neurons. Associations have been previously reported between SNPs in *MAPT* (all of which are in strong LD with our inversion-tagging SNP) and neurodegenerative disorders, including Parkinson's disease[34] and progressive supranuclear palsy[35], a rare disease whose symptoms include depression and apathy.

DISCUSSION

The discovery of genetic loci associated with SWB, depression, and neuroticism has proven elusive. Our study identified several credible associations for two main reasons. First, our analyses had greater statistical power than prior studies because ours were conducted in larger samples. Our GWAS findings—three loci

associated with SWB, two with DS, and eleven with neuroticism—support the view that GWAS can successfully identify genetic associations with highly polygenic phenotypes in sufficiently large samples[5,36]. A striking finding is that two of our identified associations are with inversion polymorphisms.

Second, our proxy-phenotype analyses further boosted power by exploiting the strong genetic overlap between our three phenotypes. These analyses identified two additional loci associated with neuroticism and two with both DS and neuroticism. Through our quasi-replication tests, we also demonstrated how studying genetically overlapping phenotypes in concert can provide evidence on the credibility of GWAS findings. Our direct replication of the two genome-wide significant associations with DS in an independent depression sample provides further confirmation of those findings (**Fig. 2b** and **Supplementary Table 15**).

We were able to assemble much larger samples than prior work in part because we combined data across heterogeneous phenotype measures. Our results reinforce the conclusions from our theoretical analysis that doing so increased our statistical power, but our strategy also has drawbacks. One is that mixing different measures may make any discovered associations more difficult to interpret. For example, since our DS phenotype is coarse and composed of varied measures, it is not clear *which* depressive symptoms are responsible for the genetic associations we found. Research studying higher quality measures of the various facets of SWB, DS, and neuroticism is a critical next step. Our results can help facilitate such work because if the variants we identify are used as candidates, studies conducted in the smaller samples in which more fine-grained phenotype measures are available can be well powered.

Another limitation of mixing different measures is that doing so may reduce the heritability of the resulting phenotype, if the measures are influenced by different genetic factors. Indeed, our estimates of SNP-based heritability[10] for our three phenotypes are quite low: 0.040 (SE = 0.002) for SWB, 0.047 (SE = 0.004) for DS, and 0.091 (SE = 0.007) for neuroticism. We correspondingly find that polygenic scores constructed from all measured SNPs explain a low fraction of variance in independent samples: ~0.9% for SWB, ~0.5% for DS, and ~0.7% for neuroticism (Online Methods). The low heritabilities imply that even when polygenic scores can be estimated using much larger samples than ours, they are unlikely to attain enough predictive power to be clinically useful.

According to our Bayesian calculations, the true explanatory power (corrected for winner's curse) of the SNP with the largest posterior R^2 is 0.003% for SWB, 0.002% for DS, and 0.011% for neuroticism (**Supplementary Table 14**). These effect sizes imply that in order to account for even a moderate share of the heritability, hundreds or (more likely) thousands of variants will be required. They also imply that our study's power to detect variants of these effect sizes was not high—for example, our statistical power to detect the lead SNP with largest posterior R^2 was only ~13%—which in turn means it is likely that there exist many variants with effect sizes comparable to our identified SNPs that evaded detection. These estimates suggest that many more loci will be found in studies with sample sizes realistically attainable in the near future.

Online Methods

This article is accompanied by a **Supplementary Note** with details on the genome-wide association analyses and follow-up analyses reported in the article.

Accession codes. Meta-analysis results can be downloaded from the SSGAC website (http://www.thessgac.org/#!data/kuzq8). For neuroticism and DS, meta-analysis results from the combined analyses are provided for all variants. For SWB, meta-analysis results for all variants are provided for the full sample excluding 23andMe, which is subject to special restrictions. For the full SWB meta-analysis, we provide results for 10,000 SNPs.

Clumping algorithm. To identify the 16 approximately independent ("lead") SNPs shown in Panel A of **Table 1**, we used the following clumping algorithm. First, the SNP with the smallest *p*-value was identified in the meta-analysis results. This SNP was designated the lead SNP of clump 1. Second, we identified all SNPs whose LD with the lead SNP exceeds $R^2 = 0.1$ and assigned them to clump 1. To generate the second clump, we removed the SNPs in clump 1 and then followed the same steps: the remaining SNP with lowest *p*-value was designated the lead SNP of clump 2, and all remaining SNPs whose LD with the lead SNP exceeds $R^2 = 0.1$ were assigned to clump 2. The process is repeated to identify further lead SNPs and their corresponding clumps until no genome-wide significant SNPs remain. The clumps define "loci" of the genome that are associated with the phenotype. For several of our other analyses (described below), we use the same clumping algorithm, albeit with lower *p*-value thresholds than genome-wide significance, to identify a set of approximately independent variants.

GWAS of SWB. Genome-wide association analyses were performed at the cohort level according to a pre-specified analysis plan. Genotyping was performed using a range of common, commercially available genotyping arrays. The analysis plan instructed cohorts to upload results imputed using the HapMap2 CEU (r22.b36) reference sample[37]. We meta-analyzed summary association statistics from 59 contributing cohorts with a combined sample size of 298,420 individuals. Before meta-analysis, a uniform set of quality-control (QC) procedures were applied to the cohort-level summary statistics, including but not limited to the EasyQC[38] protocol. All analyses were restricted to European-ancestry individuals.

We performed a sample-size-weighted meta-analysis of the cohort-level summary statistics in Metal[39]. To adjust standard errors for non-independence, we inflated them using the square root of the estimated intercept from a LD Score regression[10]. Although we consider them secondary to the SWB analyses, we also performed separate meta-analyses of PA (N = 180,281) and LS (N = 166,205) and a post hoc genome-wide analysis of SWB in cohorts with 1000G-imputed data (N = 229,883); see **Supplementary Figures 3-4** for quantile-quantile and Manhattan plots, and **Supplementary Figure 5** for LocusZoom of the two SNPs that reached genome-wide significance in the SWB analysis of 1000G-imputed data.

Detailed cohort descriptions, information about cohort-level genotyping and imputation procedures, cohort-level measures, and quality-control filters are shown in **Supplementary Tables 2-6**. **Supplementary Table 7** reports association results from the following four meta-analyses: the primary SWB analysis, the LS analysis, the PA analysis, and the post hoc SWB analysis. For each phenotype, we provide association results for the set of approximately independent SNPs that attained a *p*-value smaller than 10^{-5}. We identify these SNP using our clumping algorithm, but with the *p*-value threshold set at 10^{-5} instead of genome-wide significance.

GWAS of DS and neuroticism. Our auxiliary genome-wide association studies of DS and neuroticism were conducted in 1000G-imputed data, combining new genome-wide association analyses with publicly available summary statistics from previously published studies. We applied a similar QC protocol to that used in our primary SWB analysis. In the DS meta-analysis (N = 180,866), we weighted the UKB analysis

by sample size and the two case-control studies by effective sample size[39]. In the neuroticism meta-analysis, we performed a sample-size-weighted fixed-effects meta-analysis of the UKB data and the publicly available summary statistics from a previous GWAS of neuroticism.

Detailed cohort descriptions, information about cohort-level genotyping and imputation procedures and quality-control filters are provided in **Supplementary Tables 8-12**. See **Supplementary Figure 6** for quantile-quantile plots of the neuroticism and DS meta-analysis results. Association results for the set of approximately independent set of SNPs that attained a *p*-value smaller than 10^{-5} are supplied in **Supplementary Table 25**.

Two of our lead SNPs on chromosome 18 (rs1557341 and rs12961969) reached genome-wide significance in unconditional analyses, and their pairwise linkage disequilibrium is below our cutoff of $R^2 = 0.10$. They therefore satisfy our definition of approximate independence. We found in additional robustness analyses that the evidence that that these SNPs reflect independent genetic signals is weaker than for the remaining SNPs identified in our main analysis (**Supplementary Note**).

Population stratification. To quantify the fraction of the observed inflation of the mean test statistic that is due to bias, we used LD Score regression[10]. The estimated LD Score regression intercepts were all close to 1, suggesting no appreciable inflation of the test statistics attributable to population stratification in any of our SWB, DS, or neuroticism meta-analyses (**Supplementary Fig. 7**). For all three phenotypes, our estimates suggest that less than 2% of the observed inflation of the mean test statistic was accounted for by bias.

In our primary GWAS of SWB, we also used two family-based analyses to test for and quantify stratification biases. These analyses used within-family (WF) estimates, the coefficients from regressing the difference in phenotype across siblings on the difference in siblings' genotype (and controls). These WF estimates are not biased by population stratification because siblings share their ancestry entirely, and therefore differences in siblings' genotypes cannot be due to the siblings being from different population groups. We meta-analyzed association statistics from WF analyses conducted in four cohorts.

In the first analysis, we estimated the fraction of SNPs for which the signs of the WF estimates were concordant with the signs of the estimates obtained from a GWAS identical to our primary SWB GWAS except with the four family cohorts excluded. For the 112,884 approximately independent SNPs considered, we found a sign concordance of 50.83%, which is significantly greater than 50% ($p = 1.04 \times 10^{-8}$). Under the null hypothesis of no population stratification, the observed sign concordance matches the expected rate after winner's curse adjustment nearly perfectly, 50.83%.

The second analysis utilized the WF regression coefficient estimates (i.e., not only their signs) to estimate the amount of stratification bias. For each SNP j, let $\hat{\beta}_j$ denote the GWAS estimate, and let $\hat{\beta}_{WF,j}$ denote the WF estimate. Under the assumption that the causal effect of each SNP is the same within families as in the population, we can decompose the estimates as:

$$\hat{\beta}_j = \beta_j + s_j + U_j$$
$$\hat{\beta}_{WF,j} = \beta_j + V_j,$$

where β_j is the true underlying GWAS parameter for SNP j, s_j is the bias due to stratification (defined to be orthogonal to β_j and U_j), and U_j and V_j are the sampling variances of the estimates with $\mathrm{E}(U_j) = \mathrm{E}(V_j) = 0$. Since stratification biases are absent in WF analyses, the second equation does not have a bias term. Whenever $s_j \neq 0$, the GWAS estimate of $\hat{\beta}_j$ is biased away from the population parameter β_j. The proportion of variance in the GWAS coefficients accounted for by true genetic signals can be written as:

$$\frac{\mathrm{Var}(\beta_j)}{\mathrm{Var}(\beta_j) + \mathrm{Var}(s_j)}.$$

In **Supplementary Note**, we show that with estimates of $\hat{\beta}_j$ and $\hat{\beta}_{WF,j}$ (and their standard errors) from independent samples, it is possible to consistently estimate the above ratio. We found that with 95% confidence, between 72% and 100% of the signal in GWAS estimates is a result of true genetic effects on SWB rather than stratification.

Analyses of inversion polymorphisms. Two genome-wide significant SNPs for the neuroticism analysis are located within well-known inversion polymorphisms, on chromosomes 8 and 17. Using the genotypic data available for UKB participants, we called the inversion genotypes for UKB participants using a PCA-mixture method. For both inversions, the method clearly distinguishes 3 clusters of genotypes, corresponding to inversion genotypes (**Supplementary Fig. 8**). We validated the PCA-mixture procedure using existing methods designed to call inversion genotypes[40] (**Supplementary Table 26**).

For both inversions, we established that the inversion-tagging SNPs were always located in close proximity of the inversion region (**Fig. 3b** and **Supplementary Figs. 8-9**). **Supplementary Tables 27-28** list the twenty variants that most strongly correlate with the PCs that capture the inversion polymorphisms on chromosome 8 and 17, respectively. In additional analyses, we confirmed that the inversion is associated with neuroticism and SWB in independent cohorts (**Supplementary Tables 29-30** and **Supplementary Fig. 10**).

Proxy-phenotype analyses. In these analyses, we used a two-stage approach that has been successfully applied in other contexts[6]. In the first stage, we conducted a meta-analysis of our first-stage "proxy phenotype" and used our clumping procedure to identify the set of approximately independent SNPs at the *p*-value threshold of 10^{-4}. In the second stage, we tested SNPs identified in stage 1 (or high-LD proxies for them) for association with a second-stage phenotype in an independent (non-overlapping) sample. In our analyses, we used our primary phenotype of SWB as the proxy-phenotype. We conducted one analysis with DS as the second-stage phenotype, and one analysis with neuroticism as the second-stage phenotype. In the analyses, we omit cohorts from the first-stage or second-stage as needed to ensure that the samples in the two stages are non-overlapping. **Supplementary Table 31** lists the cohort restrictions imposed. These cohort restrictions, as well as the *p*-value threshold of 10^{-4}, were chosen before the data were analyzed on the basis of statistical power calculations.

To test for cross-phenotype enrichment, we used a non-parametric procedure that tests whether the lead SNPs are more strongly associated with the second-stage phenotype than randomly chosen sets of SNPs with a similar distribution of allele frequencies. We generated 1,000 matched SNPs for each of the Y lead/lead-proxy SNPs. We then ranked the $Y\times1000 + Y$ SNPs by *p*-value and conducted a Mann-Whitney test[41] of the null hypothesis that the *p*-value distribution of the Y lead/lead-proxy SNPs are drawn from the same distribution as the $Y\times1000$ matched SNPs.

To test the individual lead SNPs for experiment-wide significance, we examined whether any of the lead SNPs (or their high-LD proxies) are significantly associated with the second-stage phenotype at the Bonferroni-adjusted significance level of $0.05/Y$.

Genetic correlations. We used bivariate LD Score regression[10] to quantify the amount of genetic heterogeneity among the phenotypic measures pooled in each of our three separate meta-analyses. For SWB, we estimated a pairwise correlation of 0.981 (SE = 0.065) between LS and PA, 0.897 (SE = 0.017) between "WB" (our measure that combines LS and PA) and LS, and 1.031 (SE = 0.019) between PA and WB. For DS, we estimated a genetic correlation of 0.588 (SE = 0.242) between GERA and PGC, 0.972 (SE = 0.216) between GERA and UKB, and 0.797 (SE = 0.108) between UKB and PGC. Finally, we estimated a genetic

correlation of 1.11 (SE = 0.14) between the measures of neuroticism in the UKB analyses and the summary statistics from a previously published meta-analysis[4].

Using the same method, we estimated the pairwise genetic correlations between our three phenotypes of SWB, DS, and neuroticism and between each of our three phenotypes and the ten neuropsychiatric and physical health variables (for which publicly available summary statistics were available).

Bayesian credibility analyses. To evaluate the credibility of our findings, we use a standard Bayesian framework[19] in which our prior distribution for any SNP's effect is:

$$\beta \sim \begin{cases} N\left(0, \tau_j^2\right) & \text{with probability } \pi \\ 0 & \text{otherwise.} \end{cases}.$$

Here, π is the fraction of non-null SNPs, and τ_j^2 is the variance of the non-null SNPs for trait $j \in$ {SWB, DS, neuroticism}. In this framework, credibility is defined as the probability that a given SNP is non-null.

We begin with univariate analyses of the GWAS results that do not incorporate the additional information from the quasi-replication analyses of the 16 lead SNPs reported in **Table 1**. We use the three SWB-associated SNPs to illustrate our approach, but we use analogous procedures when analyzing DS and neuroticism. We calculate credibility for each value $\pi \in \{0.01, 0.02, \ldots, 0.99\}$. For each assumed value of π, we estimate τ_{SWB}^2 by maximum likelihood (**Supplementary Note**). For each SNP, we use Bayes' rule to obtain a posterior estimate of credibility for each of the assumed values of π. **Supplementary Figure 12** shows that for all considered values of π and all three SNPs, the posterior probability that the SNP is null is below 1%. Similar analyses of the DS and neuroticism SNPs show that the posterior probability never exceeds 5%.

In our joint analyses, we consider two phenotypes with genetic correlation r_g. We make the simplifying assumption that the set of null SNPs is the same for both phenotypes. The joint distribution of a SNP's effect on the two phenotypes is then given by

$$\begin{bmatrix} \beta_1 \\ \beta_2 \end{bmatrix} \sim \begin{cases} N\left(\begin{bmatrix} 0 \\ 0 \end{bmatrix}, \begin{bmatrix} \tau_1^2 & \tau_1\tau_2 r_g \\ \tau_1\tau_2 r_g & \tau_2^2 \end{bmatrix}\right) & \text{with probability } \pi \\ \begin{bmatrix} 0 \\ 0 \end{bmatrix} & \text{otherwise.} \end{cases}$$

With coefficient estimates, $\hat{\beta}_1$ and $\hat{\beta}_2$, obtained from non-overlapping samples, the variance-covariance matrix of the estimation error will be diagonal. We denote the diagonal entries of this matrix, which represent the variances of the estimation error in the two samples, by σ_1^2 and σ_2^2. This gives us the joint prior distribution

$$\begin{bmatrix} \hat{\beta}_1 \\ \hat{\beta}_2 \end{bmatrix} \sim \begin{cases} N\left(\begin{bmatrix} 0 \\ 0 \end{bmatrix}, \begin{bmatrix} \tau_1^2 & \tau_1\tau_2 r_g \\ \tau_1\tau_2 r_g & \tau_2^2 \end{bmatrix} + \begin{bmatrix} \sigma_1^2 & 0 \\ 0 & \sigma_2^2 \end{bmatrix}\right) & \text{with probability } \pi \\ N\left(\begin{bmatrix} 0 \\ 0 \end{bmatrix}, \begin{bmatrix} \sigma_1^2 & 0 \\ 0 & \sigma_2^2 \end{bmatrix}\right). & \text{otherwise.} \end{cases}$$

To select parameter values for the prior, we use the estimates of r_g reported in **Supplementary Table 1**, and we estimate the parameters π, τ_1^2, and τ_2^2 from GWAS summary statistics using a maximum likelihood procedure. For this procedure, we make the standard assumption[10,42] that the variance of a SNP's effect size is inversely proportional to the variance of its genotype, $2 \times \text{MAF} \times (1 - \text{MAF})$. (This assumption implies that rare SNPs are more likely to have large effects.)

The credibility estimates follow from applying Bayes' Rule to calculate either the probability that the SNP is non-null (an event denoted C) given only the first-stage estimate, $P\left(C \mid \hat{\beta}_1\right)$, or the probability that the SNP

is non-null conditional on the results of both the first-stage GWAS and the quasi-replication analysis, $P(C \mid \hat{\beta}_1, \hat{\beta}_2)$. Credibility estimates for our lead SNPs are in **Supplementary Table 14**.

To calculate the expected record of a replication or quasi-replication study, we assume that the SNP is non-null for both phenotypes. (This is analogous to a standard power calculation for a single phenotype, in which the SNP is assumed to be non-null.) Under this assumption, $\hat{\beta}_1$ and $\hat{\beta}_2$ are jointly normally distributed, implying that the conditional distribution of $\hat{\beta}_2$ given $\hat{\beta}_1$ is

$$(\hat{\beta}_2 \mid \hat{\beta}_1, C) \sim N\left[\frac{\tau_1 \tau_2 r_g}{\tau_1^2 + \sigma_1^2}\hat{\beta}_1, \frac{(\tau_1^2 + \sigma_1^2)(\tau_2^2 + \sigma_2^2) - \tau_1^2 \tau_2^2 r_g^2}{\tau_1^2 + \sigma_1^2}\right].$$

Using this equation, we can calculate the probability that the GWAS estimates will have concordant signs across the two phenotypes, or that the GWAS estimate of the second-stage phenotype will reach some level of significance (or other measures of replicability). These probabilities can be summed over the set of lead SNPs to generate the expected number of SNPs meeting the criterion.

To obtain effect-size estimates for a SNP that are adjusted for the winner's curse (**Supplementary Table 32**), we use the mean of the posterior distribution of the SNP's effect, conditional on the quasi-replication result and the SNP being non-null. The posterior distribution is

$$(\beta_1 \mid \hat{\beta}_1, \hat{\beta}_2, C) \sim N\left[m(\hat{\beta}_1, \hat{\beta}_2), s(\hat{\beta}_1, \hat{\beta}_2)\right]$$

where,

$$m(\hat{\beta}_1, \hat{\beta}_2) \equiv \left(\frac{\tau_1^2(\tau_2^2 + \sigma_2^2) - r_g^2 \tau_1^2 \tau_2^2}{(\tau_1^2 + \sigma_1^2)(\tau_2^2 + \sigma_2^2) - r_g^2 \tau_1^2 \tau_2^2}\right)\hat{\beta}_1 + \left(\frac{r_g \tau_1 \tau_2 \sigma_1^2}{(\tau_1^2 + \sigma_1^2)(\tau_2^2 + \sigma_2^2) - r_g^2 \tau_1^2 \tau_2^2}\right)\hat{\beta}_2$$

$$s(\hat{\beta}_1, \hat{\beta}_2) \equiv \frac{\tau_1^2 \sigma_1^2 (\tau_2^2 + \sigma_2^2 - r_g^2 \tau_2^2)}{(\tau_1^2 + \sigma_1^2)(\tau_2^2 + \sigma_2^2) - r_g^2 \tau_1^2 \tau_2^2}.$$

The winner's-curse-adjusted GWAS coefficient is simply the mean of the posterior distribution. The expected R^2 of the SNP for the first-stage phenotype is

$$E(R_1^2 \mid \hat{\beta}_1, \hat{\beta}_2) = \left[m(\hat{\beta}_1, \hat{\beta}_2)^2 + s(\hat{\beta}_1, \hat{\beta}_2)\right] \sigma_x^2 P(C \mid \hat{\beta}_1, \hat{\beta}_2),$$

where $\sigma_x^2 = 2 \times \text{MAF} \times (1 - \text{MAF})$ is the variance of the genotype under Hardy-Weinberg equilibrium. The winner's curse adjusted formulas for the second-stage phenotype are symmetric.

Lookup of DS and neuroticism-associated SNPs in an independent depression study. We partnered with the investigators of an ongoing large-scale GWAS of major depressive symptoms (N = 368,890) to follow up on the associations identified in the DS and neuroticism analyses (and reciprocally supplied them with association results for their most significant associations). The participants of the study were all European-ancestry customers of 23andMe, a personal genomics company, who responded to online survey questions about mental health. We did not request results for the SNPs identified in the SWB or proxy-phenotype analyses, since these were both conducted in samples that overlap with 23andMe's depression sample. For details on association models, quality-control filters, and the ascertainment of depression status, we refer to the companion study[21]. The p-values we report are based on standard errors that have been inflated by the square by the intercept from an LD score regression[10].

Polygenic prediction. To evaluate the predictive power of a polygenic score derived from the SWB meta-analysis results, we used two independent hold-out cohorts: the Health and Retirement Study (HRS[43]) and the Netherlands Twin Register (NTR[44,45]). To generate the weights for the polygenic score, we performed meta-analyses of the pooled SWB phenotype excluding each of the holdout cohorts, applying a minimum-

sample-size filter of 100,000 individuals. Using the summary statistics from this meta-analysis, we constructed two sets of polygenic scores: (1) LDpred polygenic scores, with weights constructed from conditional SNP effects estimated from summary statistics using information about LD structure from a reference sample[20], and (2) linear polygenic scores using the unconditional GWAS effect sizes[46]. We examined the predictive power of the score for the following outcomes: SWB, LS, PA, depression, the NEO Big Five personality traits[47], and height (the last being included as a negative control). The results from these analyses are reported in **Supplementary Table 32** and depicted in **Supplementary Figure 13**.

Enrichment of schizophrenia SNPs for association with neuroticism. In post hoc analyses, we used the test of cross-phenotype enrichment described in the "Proxy-phenotype analysis" Online Methods section to examine whether the genome-wide associations that have been reported for schizophrenia showed evidence for enriched association with neuroticism in a non-overlapping sample. Among our phenotypes, we focused on neuroticism to maximize power.

For the schizophrenia SNPs (128 of them[31]), we strongly reject the null hypothesis of no enrichment relative to a set of SNPs matched on allele frequency ($p = 6.50\times 10^{-71}$). Also, 23 of the 106 matched schizophrenia SNPs are nominally significantly associated (p-value < 0.05) with neuroticism in our sample, and 19 of these 23 SNPs have concordant signs for schizophrenia and neuroticism. For bipolar disease SNPs (five of them[48]) and anxiety disorder (two SNPs[49]), we found no evidence of enrichment.

Biological annotation. For each of SWB, DS, and neuroticism, we used stratified LD score regression[26] to test for enriched association with SNPs in (i) functional genomic regions of the genome and (i) SNPs located near histone marks in specific tissues.

For the biological annotation of the 20 SNPs in **Table 1**, we generated a list of LD partners for each of the original SNPs. A SNP was considered an LD partner for the original SNP if (i) its pairwise LD with the original SNP exceeded $R^2 = 0.6$ and (ii) it was located within 250kb of the original SNP. We also generated a list of genes residing within loci tagged by our lead SNPs (**Supplementary Table 34**).

We used the NHGRI GWAS catalog[50] to determine which of our 20 SNPs (and their LD partners) were in LD with SNPs for which genome-wide significant associations have been previously reported. Since the GWAS catalog does not always include the most recent GWAS results available, we included additional recent GWAS studies. We used the tool HaploReg[51] to identify nonsynonymous variants in LD with any of the 20 SNPs or their LD partners.

We examined whether the 20 polymorphisms in **Table 1** were associated with gene expression levels (**Supplementary Table 24**). The c*is*-eQTL associations were performed in 4,896 peripheral-blood gene expression and genome-wide SNP samples from two Dutch cohorts measured on the Affymetrix U219 platform[44,45,52]. We considered a SNP a potential *cis*-eQTL if the distance between the SNP and the midpoint of the probe set was smaller than 1Mb. The *cis*-eQTL analyses of the inversion were conducted separately. To supplement the analyses in blood, we performed eQTL lookups of our 20 SNPs in the Genotype-Tissue Expression Portal (www.GTExportal.org)[53,54]. For a given SNP, the portal provides results in various tissues for tests of association between the SNP and gene expression of genes whose transcription start site is within one Mb of the SNP. We restricted the search to the following trait-relevant tissues: hippocampus, hypothalamus, anterior cingulate cortex (BA24), putamen (basal ganglia), frontal cortex (BA9), nucleus accumbens (basal ganglia), caudate (basal ganglia), cortex, cerebellar hemisphere, cerebellum, tibial nerve, thyroid, adrenal gland, and pituitary.

Finally, using a gene co-expression database[55], we explored the predicted functions of genes co-locating with the 20 SNPs in Table 1 (**Supplementary Table 35**).

References

1. Kendler, K. S. & Myers, J. The genetic and environmental relationship between major depression and the five-factor model of personality. *Psychol. Med.* **40,** 801 (2009).
2. Weiss, A., Bates, T. C. & Luciano, M. Happiness is a personal(ity) thing: The genetics of personality and well-being in a representative sample. *Psychol. Sci.* **19,** 205–210 (2008).
3. Bartels, M., Cacioppo, J. T., van Beijsterveldt, T. C. E. M. & Boomsma, D. I. Exploring the association between well-being and psychopathology in adolescents. *Behav. Genet.* **43,** 177–190 (2013).
4. de Moor, M. H. M. *et al.* Meta-analysis of Genome-wide Association Studies for Neuroticism, and the Polygenic Association With Major Depressive Disorder. *JAMA Psychiatry* **72,** 642–650 (2015).
5. Hyman, S. Mental health: Depression needs large human-genetics studies. *Nature* **515,** 189–191 (2014).
6. Rietveld, C. A. *et al.* Common genetic variants associated with cognitive performance identified using the proxy-phenotype method. *Proc. Natl. Acad. Sci. U. S. A.* **111,** 13790–13794 (2014).
7. Kahneman, D. & Deaton, A. High income improves evaluation of life but not emotional well-being. *Proc. Natl. Acad. Sci. U. S. A.* **107,** 16489–16493 (2010).
8. Kahneman, D. & Riis, J. in *Sci. Well-Being* (Uppter, F., Baylis, N. & Keverne, B.) 285–301 (Oxford University Press, 2005).
9. Bartels, M. & Boomsma, D. I. Born to be happy? The etiology of subjective well-being. *Behav. Genet.* **39,** 605–615 (2009).
10. Bulik-Sullivan, B. K. *et al.* LD Score regression distinguishes confounding from polygenicity in genome-wide association studies. *Nat. Genet.* **47,** 291–295 (2015).
11. Yang, J. *et al.* Genomic inflation factors under polygenic inheritance. *Eur. J. Hum. Genet.* **19,** 807–812 (2011).
12. Ripke, S. *et al.* A mega-analysis of genome-wide association studies for major depressive disorder. *Mol. Psychiatry* **18,** 497–511 (2013).
13. Sudlow, C. *et al.* UK Biobank: An Open Access Resource for Identifying the Causes of a Wide Range of Complex Diseases of Middle and Old Age. *PLoS Med.* **12,** e1001779 (2015).
14. dbGaP. Resource for Genetic Epidemiology Research on Adult Health and Aging (GERA). (2015).
15. Eysenck, H. J. & Eysenck, S. B. G. *Manual of the Eysenck Personality Questionnaire*. (Hodder and Stroughton, 1975).
16. Tian, C. *et al.* Analysis and application of European genetic substructure using 300 K SNP information. *PLoS Genet.* **4,** e4 (2008).
17. Steinberg, K. M. *et al.* Structural diversity and African origin of the 17q21.31 inversion polymorphism. *Nat. Genet.* **44,** 872–80 (2012).
18. Smith, D. J. *et al.* Genome-wide analysis of over 106,000 individuals identifies 9 neuroticism-associated loci. *bioRxiv* (2015). doi:http://dx.doi.org/10.1101/032417
19. Meuwissen, T. H., Hayes, B. J. & Goddard, M. E. Prediction of total genetic value using genome-wide dense marker maps. *Genetics* **157,** 1819–1829 (2001).
20. Vilhjálmsson, B. J. *et al.* Modeling Linkage Disequilibrium Increases Accuracy of Polygenic Risk Scores. *Am. J. Hum. Genet.* **97,** 576–592 (2015).
21. Hyde, C. L. et al. Common genetic variants associated with major depressive disorder among individuals of European descent. *Nat. Genet.*
22. Roberts, R. E., Kaplan, G. a, Shema, S. J. & Strawbridge, W. J. Are the obese at greater risk for

depression? *Am. J. Epidemiol.* **152,** 163–170 (2000).

23. Glassman, A. J. *et al.* Smoking, smoking cessation, and major depression. *J. Am. Med. Assoc.* **264,** 1546–1549 (1990).

24. Shahab, L. & West, R. Differences in happiness between smokers, ex-smokers and never smokers: Cross-sectional findings from a national household survey. *Drug Alcohol Depend.* **121,** 38–44 (2012).

25. Rugulies, R. Depression as a predictor for coronary heart disease: A review and meta-analysis. *Am. J. Prev. Med.* **23,** 51–61 (2002).

26. Finucane, H. K. *et al.* Partitioning heritability by functional category using GWAS summary statistics. *Nat. Genet.* **47,** 1228–1235 (2015).

27. Wood, A. R. *et al.* Defining the role of common variation in the genomic and biological architecture of adult human height. *Nat. Genet.* **46,** 1173–1186 (2014).

28. Stetler, C. & Miller, G. E. Depression and hypothalamic-pituitary-adrenal activation: a quantitative summary of four decades of research. *Psychosom. Med.* **73,** 114–26

29. Seeman, P. Dopamine D2 receptors as treatment targets in schizophrenia. *Clin. Schizophr. Relat. Psychoses* **4,** 56–73 (2010).

30. Vallone, D., Picetti, R. & Borrelli, E. Structure and function of dopamine receptors. *Neurosci. Biobehav. Rev.* **24,** 125–132 (2000).

31. Ripke, S. *et al.* Biological insights from 108 schizophrenia-associated genetic loci. *Nature* **511,** 421–427 (2014).

32. Shungin, D. *et al.* New genetic loci link adipose and insulin biology to body fat distribution. *Nature* **518,** 187–196 (2015).

33. Kathiresan, S. *et al.* Common variants at 30 loci contribute to polygenic dyslipidemia. *Nat. Genet.* **41,** 56–65 (2009).

34. Spencer, C. C. A. *et al.* Dissection of the genetics of Parkinson's disease identifies an additional association 5' of SNCA and multiple associated haplotypes at 17q21. *Hum. Mol. Genet.* **20,** 345–353 (2011).

35. Höglinger, G. U. *et al.* Identification of common variants influencing risk of the tauopathy progressive supranuclear palsy. *Nat. Genet.* **43,** 699–705 (2011).

36. Sullivan, P. F. Don't give up on GWAS. *Mol. Psychiatry* **17,** 2–3 (2012).

37. The International HapMap Consortium. A second generation human haplotype map of over 3.1 million SNPs. *Nature* **449,** 851–861 (2007).

38. Winkler, T. W. *et al.* Quality control and conduct of genome-wide association meta-analyses. *Nat. Protoc.* **9,** 1192–1212 (2014).

39. Willer, C. J., Li, Y. & Abecasis, G. R. METAL: Fast and efficient meta-analysis of genomewide association scans. *Bioinformatics* **26,** 2190–2191 (2010).

40. Cáceres, A. & González, J. R. Following the footprints of polymorphic inversions on SNP data: from detection to association tests. *Nucleic Acids Res.* **43,** e53 (2015).

41. Nachar, N. The Mann-Whitney U: A test for assessing whether two independent samples come from the same distribution. *Quant. Methods Psychol.* **4,** 13–20 (2008).

42. Yang, J. *et al.* Common SNPs explain a large proportion of heritability for human height. *Nat. Genet.* **42,** 565–569 (2010).

43. Sonnega, A. *et al.* Cohort Profile: the Health and Retirement Study (HRS). *Int. J. Epidemiol.* **43,** 576–585 (2014).

44. Willemsen, G. *et al.* The Adult Netherlands Twin Register: Twenty-Five Years of Survey and

Biological Data Collection. *Twin Res. Hum. Genet.* **16,** 271–281 (2013).

45. van Beijsterveldt, C. E. M. *et al.* The Young Netherlands Twin Register (YNTR): Longitudinal Twin and Family Studies in Over 70,000 Children. *Twin Res. Hum. Genet.* **16,** 252–267 (2013).
46. Purcell, S. M. *et al.* Common polygenic variation contributes to risk of schizophrenia and bipolar disorder. *Nature* **460,** 748–752 (2009).
47. Costa, P. T. & McCrae, R. R. The five-factor model of personality and its relevance to personality disorders. *J. Pers. Disord.* **6,** 343–359 (1992).
48. Mühleisen, T. W. *et al.* Genome-wide association study reveals two new risk loci for bipolar disorder. *Nat. Commun.* **5,** 3339 (2014).
49. Otowa, T. *et al.* Meta-analysis of genome-wide association studies of anxiety disorders. *Mol. Psychiatry* 1–9 (2016). doi:10.1038/mp.2015.197
50. Welter, D. *et al.* The NHGRI GWAS Catalog, a curated resource of SNP-trait associations. *Nucleic Acids Res.* **42,** D1001–D1006 (2014).
51. Ward, L. D. & Kellis, M. HaploReg: A resource for exploring chromatin states, conservation, and regulatory motif alterations within sets of genetically linked variants. *Nucleic Acids Res.* **40,** (2012).
52. Penninx, B. W. J. H. *et al.* The Netherlands Study of Depression and Anxiety (NESDA): rationale, objectives and methods. *Int. J. Methods Psychiatr. Res.* **17,** 121–140 (2008).
53. The GTEx Consortium. The Genotype-Tissue Expression (GTEx) project. *Nat. Genet.* **45,** 580–5 (2013).
54. Ardlie, K. G. *et al.* The Genotype-Tissue Expression (GTEx) pilot analysis: Multitissue gene regulation in humans. *Science* **348,** 648–660 (2015).
55. Fehrmann, R. S. N. *et al.* Gene expression analysis identifies global gene dosage sensitivity in cancer. *Nat. Genet.* **47,** 115–25 (2015).

Acknowledgements: This research was carried out under the auspices of the Social Science Genetic Association Consortium (SSGAC). The SSGAC seeks to facilitate studies that investigate the influence of genes on human behavior, well-being, and social-scientific outcomes using large genome-wide association study meta-analyses. The SSGAC also provides opportunities for replication and promotes the collection of accurately measured, harmonized phenotypes across cohorts. The SSGAC operates as a working group within the CHARGE consortium. This research has also been conducted using the UK Biobank Resource. The study was supported by funding from the U.S. National Science Foundation (EAGER: "Workshop for the Formation of a Social Science Genetic Association Consortium"), a supplementary grant from the National Institute of Health Office of Behavioral and Social Science Research, the Ragnar Söderberg Foundation (E9/11), the Swedish Research Council (421-2013-1061), The Jan Wallander and Tom Hedelius Foundation, an ERC Consolidator Grant (647648 EdGe), the Pershing Square Fund of the Foundations of Human Behavior, and the NIA/NIH through grants P01-AG005842, P01-AG005842-20S2, P30-AG012810, and T32-AG000186-23 to NBER and R01-AG042568-02 to the University of Southern California. We are grateful to Peter M. Visscher for advice, support, and feedback. We thank Samantha Cunningham and Nishanth Galla for research assistance. A full list of acknowledgments is provided in the **Supplementary Note**. The authors declare no competing financial interests.

Authors:

Aysu Okbay[1,2,3,*], Bart M.L. Baselmans[4,5,*], Jan-Emmanuel De Neve[6,*], Patrick Turley[7,*], Michel G. Nivard[4,*], Mark Alan Fontana[8,*], S. Fleur W. Meddens[9,3,10*], Richard Karlsson Linnér[9,3,10*], Cornelius A. Rietveld[1,2,3,*], Jaime Derringer[11], Jacob Gratten[129], James J. Lee[12], Jimmy Z. Liu[135], Ronald de Vlaming[1,2,3], Tarunveer S. Ahluwalia[13,14,15], Jadwiga Buchwald[16], Alana Cavadino[17,99], Alexis C. Frazier-Wood[18], Nicholas A. Furlotte[19], Victoria Garfield[122], Marie Henrike Geisel[44], Juan R. Gonzalez[20,21,22], Saskia Haitjema[113], Robert Karlsson[23], Sander W. van der Laan[113], Karl-Heinz Ladwig[24], Jari Lahti[25,26,79], Sven J. van der Lee[2], Penelope A. Lind[27], Tian Liu[28,29] , Lindsay Matteson[12], Evelin Mihailov[30], Michael B. Miller[12], Camelia C. Minica[4], Ilja M. Nolte[34], Dennis Mook-Kanamori[31,32,33], Peter J. van der Most[34], Christopher Oldmeadow[35,36], Yong Qian[37], Olli Raitakari[38,39], Rajesh Rawal[40], Anu Realo[41], Rico Rueedi[42,43], Börge Schmidt[44], Albert V. Smith[45,46], Evie Stergiakouli[47], Toshiko Tanaka[61], Kent Taylor[48], Juho Wedenoja[16], Juergen Wellmann[49], Harm-Jan Westra[50], Sara M. Willems[2], Wei Zhao[51], Behrooz Z. Alizadeh[34,52], Najaf Amin[2], Andrew Bakshi[129], Patricia A. Boyle[53], Samantha Cherney[54], Simon Cox[55,56], Gail Davies[55,56], Oliver S.P. Davis[47], Jun Ding[37], Nese Direk[2], Peter Eibich[57,58], Rebecca T. Emeny[24], Ghazaleh Fatemifar[59], Jessica D. Faul[60], Luigi Ferrucci[61], Andreas J. Forstner[62,63], Christian Gieger[40], Richa Gupta[16], Tamara B. Harris[64], Juliette M. Harris[80], Elizabeth G. Holliday[35,36], Jouke-Jan Hottenga[4,5], Philip L. De Jager[65,66,67], Marika A. Kaakinen[68,71], Eero Kajantie[69,70], Ville Karhunen[71], Ivana Kolcic[116], Meena Kumari[72], Lenore J. Launer[73], LifeLines Cohort Study[74], Ruifang Li-Gao[31], Marisa Loitfelder[119], Anu Loukola[16], Pedro Marques-Vidal[75], Grant W. Montgomery[76], Miriam A. Mosing[77], Lavinia Paternoster[47], Alison Pattie[56], Katja E. Petrovic[78], Laura Pulkki-Råback[25,79], Lydia Quaye[80], Katri Räikkönen[25], Igor Rudan[81], Rodney J. Scott[82,36], Jennifer A. Smith[51], Angelina R. Sutin[123,61], Maciej Trzaskowski[115,129], Anna E. Vinkhuyzen[129], Lei Yu[83], Delilah Zabaneh[115], John R. Attia[35,36], David A. Bennett[83], Klaus Berger[49], Lars Bertram[84,85], Dorret I. Boomsma[4,5,138], Ute Bultmann[86], Shun-Chiao Chang[87], Francesco Cucca[88], Ian J. Deary[55,56], Cornelia M. van Duijn[2], Johan G. Eriksson[89,90,91] , Lude Franke[92], Eco J.C. de Geus[4,5,138], Patrick J.F. Groenen[3,93], Vilmundur Gudnason[45,46], Torben Hansen[14], Catharine A. Hartman[112], Claire M.A. Haworth[47], Caroline Hayward[94,95], Andrew C. Heath[96], David A. Hinds[19], Elina Hyppönen[97,98,99], William G. Iacono[12], Marjo-Riitta Järvelin[100,101,71,102], Karl-Heinz Jöckel[44], Jaakko Kaprio[16,103,104], Sharon L.R. Kardia[51], Liisa Keltikangas-Järvinen[25], Peter Kraft[105], Laura Kubzansky[106], Terho Lehtimäki[107,108], Patrik K.E. Magnusson[23], Nicholas G. Martin[109], Matt McGue[12], Andres Metspalu[30,110], Melinda Mills[111], Renée de Mutsert[31], Albertine J. Oldehinkel[112], Gerard Pasterkamp[113,114], Nancy L. Pedersen[23], Robert Plomin[115], Ozren Polasek[116], Christine Power[99], Stephen S. Rich[117], Frits R. Rosendaal[31], Hester M. den Ruijter[113], David Schlessinger[37], Helena Schmidt[118,119], Rauli Svento[120], Reinhold Schmidt[119], Harold Snieder[34], Thorkild I.A. Sørensen[14,47,121], Tim D. Spector[80], Andrew Steptoe[122], Antonio Terracciano[123,61] , A. Roy Thurik[1,3,124,125], Nicholas J. Timpson[47], Henning Tiemeier[2,126,127], André G. Uitterlinden[2,3,128], Peter Vollenweider[75], Gert Wagner[57], David R. Weir[60], Jian Yang[129,130], Dalton C. Conley[131,132], George Davey Smith[47], Albert Hofman[2], Magnus Johannesson[133], David I. Laibson[7], Sarah E. Medland[27], Michelle N. Meyer[134], Joseph K. Pickrell[135,136], Tõnu Esko[30], Robert F. Krueger[12,#], Jonathan P. Beauchamp[7,#], Philipp D. Koellinger[9,3,10,#], Daniel J. Benjamin[137,#], Meike Bartels[4,5,138,#], David Cesarini[139,140,#]

1. Department of Applied Economics, Erasmus School of Economics, Erasmus University Rotterdam, 3062 PA, Rotterdam, The Netherlands
2. Department of Epidemiology, Erasmus Medical Center, Rotterdam, 3015 GE, The Netherlands
3. Erasmus University Rotterdam Institute for Behavior and Biology, Rotterdam, 3062 PA, The Netherlands
4. Department of Biological Psychology, Vrije Universiteit, Amsterdam, 1081 BT, The Netherlands
5. EMGO+ Institute for Health and Care Research, Amsterdam, 1081 BT, The Netherlands
6. Saïd Business School, University of Oxford, Oxford, OX1 1HP, UK
7. Department of Economics, Harvard University, Cambridge, MA 02138, USA
8. Center for Economic and Social Research, University of Southern California, Los Angeles, CA 90089-3332, USA
9. Department of Complex Trait Genetics, VU University, Center for Neurogenomics and Cognitive Research, Amsterdam, 1081 HV, The Netherlands
10. Amsterdam Business School, University of Amsterdam, Amsterdam, 1018 TV, The Netherlands
11. Psychology, University of Illinois, IL 61820, Champaign, USA
12. Department of Psychology, University of Minnesota Twin Cities, Minneapolis, MN 55455, USA
13. COPSAC, Copenhagen Prospective Studies on Asthma in Childhood, Herlev and Gentofte Hospital, University of Copenhagen, Copenhagen, 2820, Denmark
14. The Novo Nordisk Foundation Center for Basic Metabolic Research, Section of Metabolic Genetics, University of Copenhagen, Faculty of Health and Medical Sciences, Copenhagen, 2100, Denmark
15. Steno Diabetes Center, Gentofte, 2820, Denmark
16. Department of Public Health, University of Helsinki, Helsinki, FI-00014, Finland
17. Centre for Environmental and Preventive Medicine, Wolfson Institute of Preventive Medicine, Queen Mary University of London, London EC1M 6BQ, UK
18. USDA/ARS Children's Nutrition Research Center, Baylor College of Medicine, Houston, TX 77030, USA
19. 23andMe, Inc., Mountain View, CA 94041, USA
20. Centre for Research in Environmental Epidemiology, Institute for Global Health, Barcelona, Spain
21. Universitat Pompeu Fabra, Barcelona, Spain
22. CIBER Epidemiología y Salud Pública, Barcelona, Spain
23. Department of Medical Epidemiology and Biostatistics, Karolinska Institutet, Stockholm, 171 77, Sweden
24. Institute of Epidemiology II, Helmholtz Zentrum München, German Research Center for Environmental Health, Neuherberg, 85764, Germany
25. IBS, Unit of Personality, Work and Health, Institute of Behavioural Sciences, P.O. Box 9, 00014 University of Helsinki, Finland
26. Folkhälsan Research Centre, Helsingfors, FI-00014, Finland
27. Quantitative Genetics, QIMR Berghofer Medical Research Institute, Brisbane, QLD 4029, Australia
28. Lifespan Psychology, Max Planck Institute for Human Development, Berlin, 14195, Germany
29. Max Planck Institute for Molecular Genetics, Department of Vertebrate Genomics, Berlin, 14195, Germany

30. Estonian Genome Center, University of Tartu, Tartu, 51010, Estonia
31. Clinical Epidemiology, Leiden University Medical Center, Leiden, 2300 RC, The Netherlands
32. Public Health and Primary Care, Leiden University Medical Center, Leiden, 2300 RC, The Netherlands
33. BESC, King Faisal Specialist Hospital and Research Centre, Riyadh 11211, Saudi Arabia
34. Department of Epidemiology, University of Groningen, University Medical Center Groningen, Groningen, 9700 RB, The Netherlands
35. Public Health Stream, Hunter Medical Research Institute, New Lambton, NSW 2305, Australia
36. Faculty of Health and Medicine, University of Newcastle, Newcastle, NSW 2300, Australia
37. Laboratory of Genetics, National Institute on Aging, Baltimore, MD 21224, USA
38. Research Centre of Applied and Preventive Cardiovascular Medicine, University of Turku, Turku, 20521, Finland
39. Department of Clinical Physiology, Turku University Hospital, Turku 20520, Finland
40. Research Unit of Molecular Epidemiology, Helmholtz Zentrum München, German Research Center for Environmental Health, Neuherberg, 85764, Germany
41. Department of Psychology, University of Tartu, Tartu 50409, Estonia
42. Department of Medical Genetics, University of Lausanne, Lausanne, 1005, Switzerland
43. Swiss Institute of Bioinformatics, Lausanne, 1015, Switzerland
44. Institute for Medical Informatics, Biometry and Epidemiology, University Hospital of Essen, Essen, 45147, Germany
45. Icelandic Heart Association, Kopavogur, 201, Iceland
46. Faculty of Medicine, University of Iceland, Reykjavik, 101, Iceland
47. MRC Integrative Epidemiology Unit, University of Bristol, Bristol, BS8 2BN, UK
48. Los Angeles Biomedical Research Institute and Department of Pediatrics, Harbor-UCLA, Torrence, 90505 CA, USA
49. Institute of Epidemiology and Social Medicine, University of Muenster, Muenster, 48149, Germany
50. Divisions of Genetics and Rheumatology, Department of Medicine, Brigham and Women's Hospital, Harvard Medical School, Boston, MA 02115, USA
51. Department of Epidemiology, University of Michigan, Ann Arbor, MI 48104, USA
52. Department of Gastroenterology and Hepatology, University of Groningen, University Medical Center Groningen, Groningen, 9713 GZ, The Netherlands
53. Department of Behavioral Sciences, Rush University Medical Center, Chicago, IL 60612, USA
54. RAND Corporation, Santa Monica, CA 9041-3208, USA
55. Centre for Cognitive Ageing and Cognitive Epidemiology, University of Edinburgh, Edinburgh, EH8 9JZ, UK
56. Department of Psychology, University of Edinburgh, Edinburgh, EH8 9JZ, UK
57. German Socio-Economic Panel Study, DIW Berlin, Berlin, 10117, Germany
58. Health Economics Research Centre, Nuffield Department of Population Health, University of Oxford, Oxford, OX3 7LF, UK
59. The Farr Institute of Health Informatics, University College London, London NW1 2DA, UK

60. Survey Research Center, Institute for Social Research, University of Michigan, Ann Arbor, MI 48104, USA

61. National Institute on Aging, National Institutes of Health, Baltimore, MD 21224, USA

62. Institute of Human Genetics, University of Bonn, Bonn, 53127, Germany

63. Department of Genomics, Life and Brain Center, University of Bonn, Bonn, 53127, Germany

64. Laboratory of Epidemiology, Demography, National Institute on Aging, National Institutes of Health, Bethesda, MD 20892-9205, United States

65. Program in Translational NeuroPsychiatric Genomics, Departments of Neurology & Psychiatry, Brigham and Women's Hospital, Boston, MA 02115, USA

66. Harvard Medical School, Boston, MA 02115, USA

67. Program in Medical and Population Genetics, Broad Institute of Harvard and Massachusetts Institute of Technology, Cambridge, MA, 02142, USA

68. Department of Genomics of Common Disease, Imperial College London, London, W12 0NN, UK

69. Department of Pediatrics, University of Helsinki, Helsinki, Finland

70. National Institute for Health and Welfare, Helsinki, Finland

71. Center for Life Course Health Research, University of Oulu; Oulu University Hospital, Oulu, Finland

72. Institute for Social & Economic Research, University of Essex, Wivenhoe Park CO4 3SQ, UK

73. Neuroepidemiology Section, National Institute on Aging, National Institutes of Health, Bethesda, MD 20892-9205, USA

74. LifeLines Cohort Study, University of Groningen, University Medical Center Groningen, Groningen, 9713 BZ, The Netherlands

75. Department of Internal Medicine, Internal Medicine, Lausanne University Hospital (CHUV), Lausanne, 1011, Switzerland

76. Molecular Epidemiology, QIMR Berghofer Medical Research Institute, Brisbane, QLD 4029, Australia

77. Department of Neuroscience, Karolinska Institutet, Retzius Väg 8 171 65 Stockholm, Sweden

78. Centre for Genomic and Experimental Medicine, Institute of Genetics and Molecular Medicine, University of Edinburgh, Edinburgh, EH4 2XU, UK

79. Helsinki Collegium for Advanced Studies, University of Helsinki, Helsinki 00014, Finland

80. Department of Twin Research and Genetic Epidemiology, King's College London, London, SE1 7EH, UK

81. Centre for Global Health Research, The Usher Institute for Population Health Sciences and Informatics, University of Edinburgh, Edinburgh, EH8 9AG, UK

82. Information Based Medicine Stream, Hunter Medical Research Institute, New Lambton, NSW 2305, Australia

83. Department of Neurological Sciences, Rush University Medical Center, Chicago, IL 60612, USA

84. Lübeck Interdisciplinary platform for Genome Analytics (LIGA), Institutes of Neurogenetics and Integrative & Experimental Genomics, University of Lübeck, Lübeck, 23562, Germany

85. Neuroepidemiology and Ageing Research Unit, School of Public Health, Faculty of Medicine, Imperial College, London SW7 2AZ, UK

86. Department of Health Sciences, Community & Occupational Medicine, University of Groningen, University Medical Center Groningen, Groningen, 9713 AV, The Netherlands

87. Channing Division of Network Medicine, Department of Medicine, Brigham and Women's Hospital, Boston, MA 02115, USA

88. Istituto di Ricerca Genetica e Biomedica (IRGB), Consiglio Nazionale delle Ricerche, c/o Cittadella Universitaria di Monserrato, Monserrato, Cagliari, 9042, Italy

89. Department of General Practice and Primary Health Care, University of Helsinki, Helsinki, 00014, Finland

90. Department of Chronic Disease Prevention, National Institute for Health and Welfare, Helsinki, Finland

91. Unit of General Practice, University Central Hospital, Helsinki, Finland

92. Department of Genetics, University Medical Center Groningen, University of Groningen, Groningen. 9700 RB, The Netherlands

93. Econometric Institute, Erasmus School of Economics, Erasmus University Rotterdam, Rotterdam, 3062 PA, The Netherlands

94. MRC Human Genetics Unit, Institute of Genetics and Molecular Medicine, University of Edinburgh, Edinburgh, EH4 2XU, UK

95. Generation Scotland, Centre for Genomics and Experimental Medicine, Institute of Genetics and Molecular Medicine, University of Edinburgh, Edinburgh, EH4 2XU, UK

96. Department of Psychiatry, Washington University School of Medicine, St. Louis, MO 63110, USA

97. Centre for Population Health Research, School of Health Sciences and Sansom Institute, University of South Australia, SA5000, Adelaide, Australia

98. South Australian Health and Medical Research Institute, Adelaide, SA5000, Australia

99. Population, Policy and Practice, UCL Institute of Child Health, London, WC1N 1EH, UK

100. Department of Epidemiology and Biostatistics, MRC Health Protection Agency (HPE) Centre for Environment and Health, School of Public Health, Imperial College London, UK

101. Biocenter Oulu, University of Oulu, Oulu, Finland

102. Unit of Primary Care, Oulu University Hospital, Oulu, Finland

103. Institute for Molecular Medicine Finland (FIMM), University of Helsinki, Helsinki, 00014, Finland

104. Department for Health, THL-National Institute for Health and Welfare, Helsinki, FI-00271, Finland

105. Department of Epidemiology and Biostatistics, Harvard T.H. Chan School of Public Health, Boston, MA 02115, USA

106. Department of Social and Behavioral Sciences, Harvard T.H. Chan School of Public Health, Boston, MA 02115, USA

107. Fimlab Laboratories, Tampere, 33520, Finland

108. Department of Clinical Chemistry, University of Tampere, School of Medicine, Tampere, 33014, Finland

109. Genetic Epidemiology, QIMR Berghofer Medical Research Institute, Brisbane, QLD 4029, Australia

110. Institute of Molecular and Cell Biology, University of Tartu, Tartu, 51010, Estonia

111. Department of Sociology, University of Oxford, Oxford OX1 3UQ, UK

112. Department of Psychiatry, University of Groningen, University Medical Center Groningen, 9700 RB, Groningen, The Netherlands

113. Laboratory of Experimental Cardiology, University Medical Center Utrecht, Utrecht, 3584 CX, The Netherlands

114. Laboratory of Clinical Chemistry and Hematology, Division Laboratories and Pharmacy, University Medical Center Utrecht, Utrecht, 3584 CX, The Netherlands

115. Social, Genetic and Developmental Psychiatry Centre, King's College London, De Crespigny Park SE5 8AF, UK

116. Department of Public Health, Faculty of Medicine, University of Split, Croatia, Split 21000, Croatia

117. Department of Public Health Sciences, University of Virginia, Charlottesville, VA 22904, USA

118. Research Unit for Genetic Epidemiology, Institute of Molecular Biology and Biochemistry, Center of Molecular Medicine, General Hospital and Medical University, Graz, Graz, 8010, Austria

119. Department of Neurology, General Hospital and Medical University Graz, Graz, 8036, Austria

120. Department of Economics, Oulu Business School, Oulu, Finland

121. Institute of Preventive Medicine, Bispebjerg and Frederiksberg Hospitals, The Capital Region, Frederiksberg, 2000, Denmark

122. Department of Epidemiology & Public Health, University College London, London WC1E 7HB, UK

123. Department of Geriatrics, Florida State University College of Medicine, Tallahassee, FL 32306, USA

124. Montpellier Business School, Montpellier, 34080, France

125. Panteia, Zoetermeer, 2715 CA, The Netherlands

126. Department of Psychiatry, Erasmus Medical Center, Rotterdam, 3015 GE, The Netherlands

127. Department of Child and Adolescent Psychiatry, Erasmus Medical Center, Rotterdam, 3015 GE, The Netherlands

128. Department of Internal Medicine, Erasmus Medical Center, Rotterdam, 3015 GE, The Netherlands

129. Queensland Brain Institute, The University of Queensland, Brisbane, QLD 4072, Australia

130. The University of Queensland Diamantina Institute, The Translation Research Institute, Brisbane, QLD 4102, Australia

131. Department of Sociology, New York University, New York, NY 10012, USA

132. School of Medicine, New York University, NY 10016, New York, USA

133. Department of Economics, Stockholm School of Economics, Stockholm, 113 83, Sweden

134. Bioethics Program, Union Graduate College - Icahn School of Medicine at Mount Sinai, Schenectady, NY 12308, USA

135. New York Genome Center, New York, NY 10013, USA

136. Department of Biological Sciences, Columbia University, 600 Fairchild Center, New York, NY 10027, USA

137. Center for Economic and Social Research, University of Southern California, Los Angeles, CA 90089-3332, USA

138. Neuroscience Campus Amsterdam, Amsterdam, The Netherlands

139. Department of Economics, New York University, New York, NY 10012, USA

140. Research Institute for Industrial Economics, Stockholm, 10215, Sweden

Fig. 1. Manhattan plots. (a) Subjective well-being (N = 298,420), **(b)** Depressive symptoms (N = 180,866), **(c)** Neuroticism (N = 170,911). The x-axis is chromosomal position, and the y-axis is the significance on a $-\log_{10}$ scale. The upper dashed line marks the threshold for genome-wide significance ($p = 5\times10^{-8}$); the lower line marks the threshold for nominal significance ($p = 10^{-5}$). Each approximately independent genome-wide significant association ("lead SNP") is marked by ×. Each lead SNP is the lowest p-value SNP within the locus, as defined by our clumping algorithm (Online Methods).

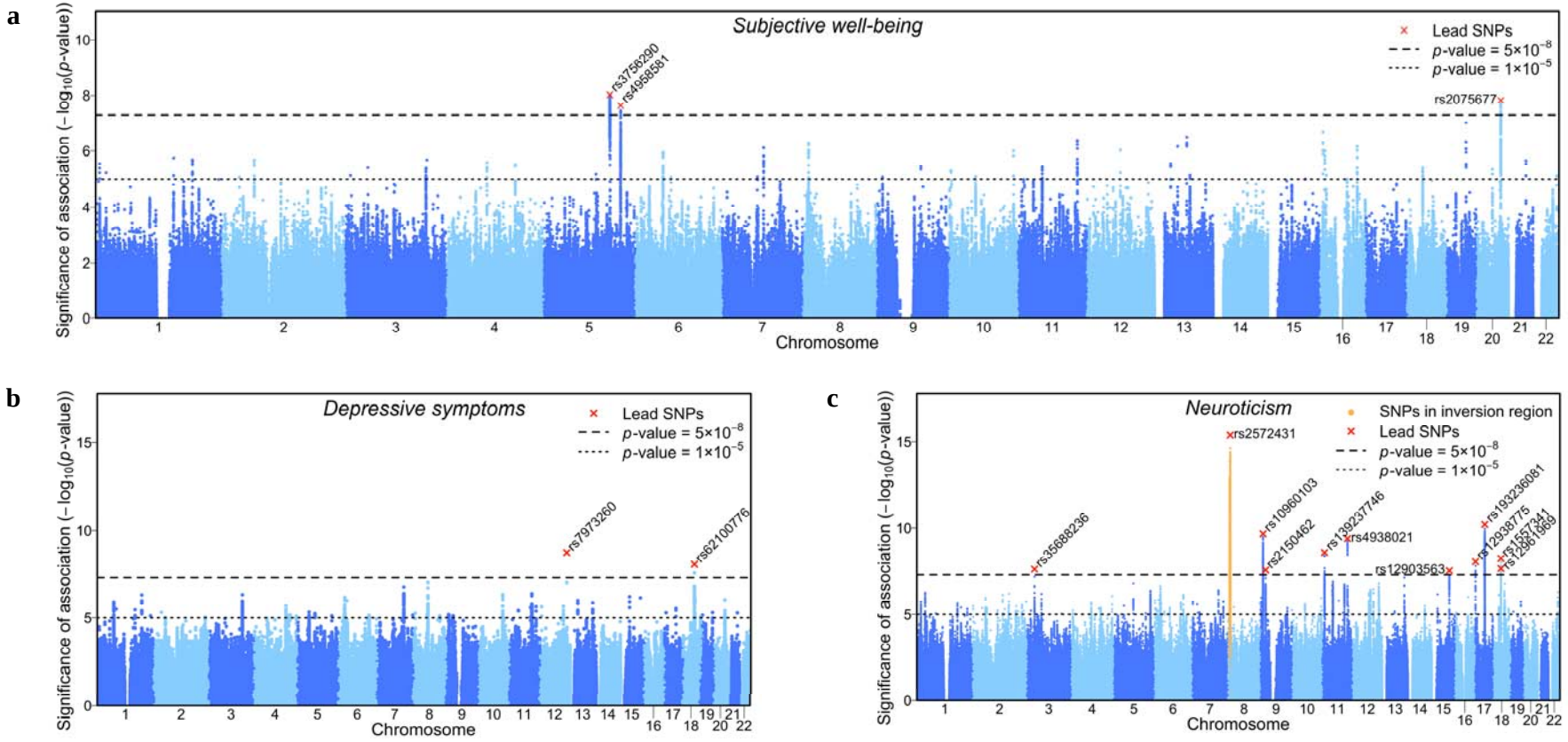

Fig. 2. Genetic correlations. The correlations are estimated using bivariate LD Score (LDSC) regression. (**a**) Genetic correlations between SWB, DS, and neuroticism ("our three phenotypes"), as well as between our three phenotypes and height. (**b**) Genetic correlations between our three phenotypes and selected neuropsychiatric phenotypes. (**c**) Genetic correlations between our three phenotypes and selected physical health phenotypes. In (**b**) and (**c**), we report the negative of the estimated correlation with DS and neuroticism (but not SWB).

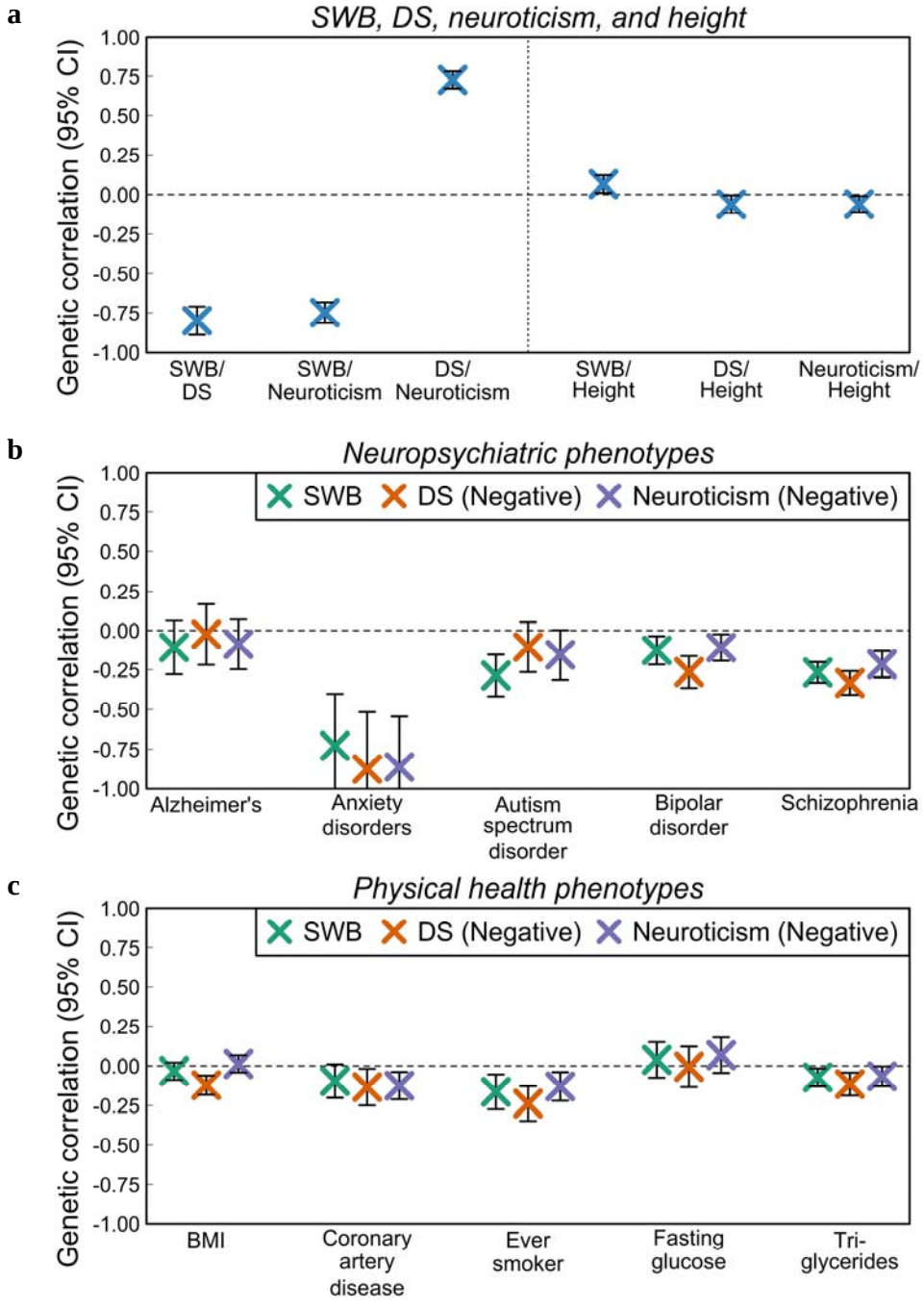

Fig. 3. Quasi-replication and lookup of lead SNPs. In quasi-replication analyses, we examined whether (**a**) lead SNPs identified in the SWB meta-analyses are associated with DS or neuroticism, (**b**) lead SNPs identified in the analyses of DS are associated with SWB, and (**c**) lead SNPs identified in the analyses of neuroticism are associated with SWB. The quasi-replication sample is always restricted to non-overlapping cohorts. In a separate lookup exercise, we examined whether lead SNPs for DS and neuroticism are associated with depression in an independent sample of 23andMe customers (N = 368,890). The results from this lookup are depicted as green crosses in (**b**) and (**c**). Bars represent 95% CIs (not adjusted for multiple testing). For interpretational ease, we choose the reference allele so that positive coefficients imply that the estimated effect is in the predicted direction. Listed below each lead SNP is the nearest gene.

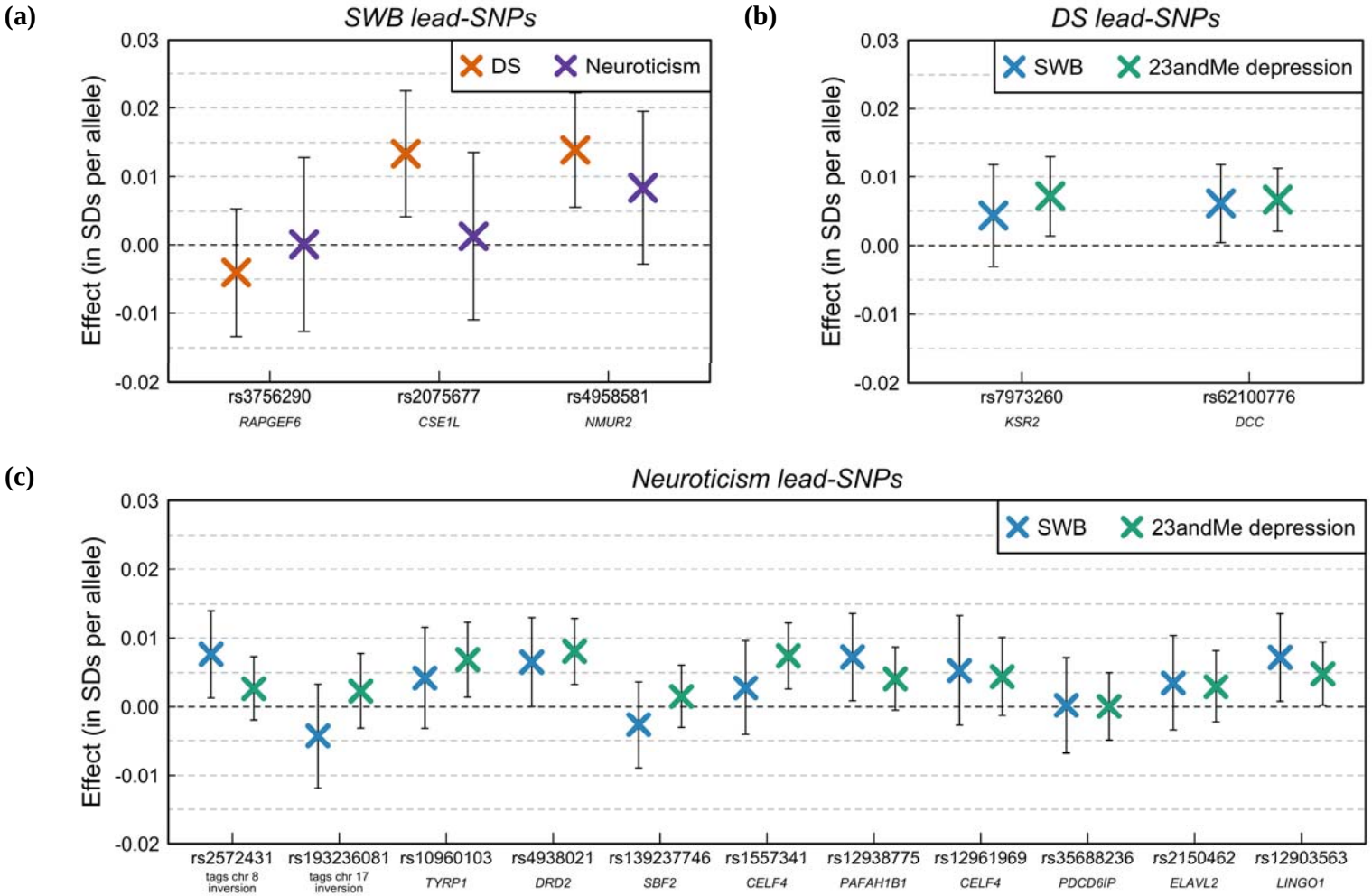

Fig. 4. Results from selected biological analyses. (**a**) Estimates of the expected increase in the phenotypic variance accounted for by a SNP due to the SNP's being in a given category (τ_c), divided by the LD Score heritability of the phenotype (h^2). Each estimate of τ_c comes from a separate stratified LD Score regression, controlling for the 52 functional annotation categories in the "baseline model." The bars represent 95% CIs (not adjusted for multiple testing). To benchmark the estimates, we compare them to those obtained from a recent study of height[27]. (**b**) Inversion polymorphism on chromosome 8 and the 7 genes for which the inversion is a significant *cis*-eQTL at FDR < 0.05. The upper half of the figure shows the Manhattan plot for neuroticism for the inversion and surrounding regions. The bottom half shows the squared correlation between the SNPs and the principal component that captures the inversion. The inlay plots the relationship, for each SNP in the inversion region, between the SNP's significance and its squared correlation with the principal component that captures the inversion.

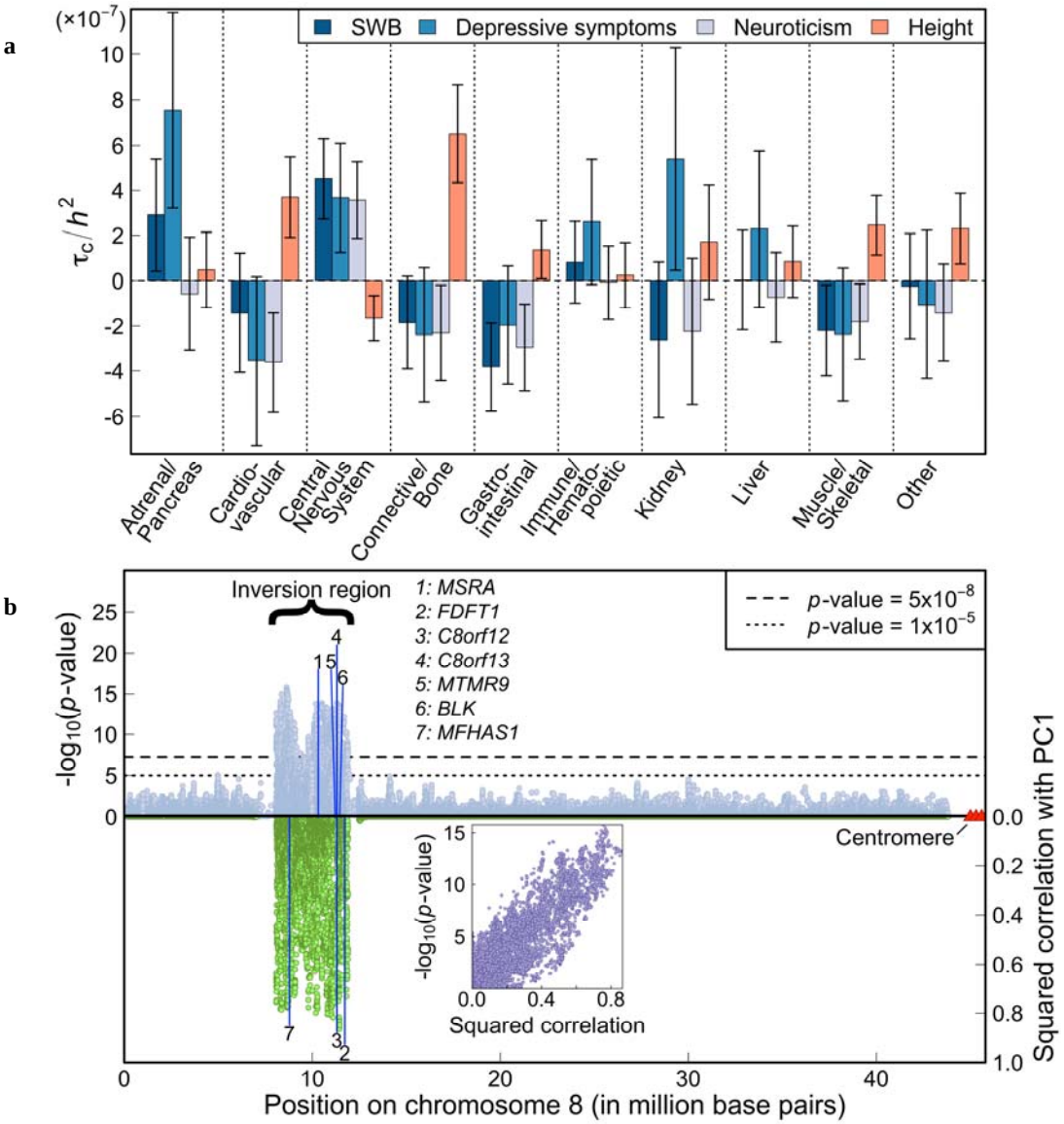

Table 1. Summary of polymorphisms identified across analyses. EA: effect allele. EAF: effect allele frequency. All effect sizes are reported in units of SDs per allele. "Quasi-Repl.": phenotypes for which SNP was found to be nominally associated in quasi-replication analyses conducted in independent samples. *significant at the 5%-level, **significant at the 1%-level, ***significant at the 0.1%-level. #inversion-tagging polymorphism on chromosome 8. ##inversion-tagging polymorphism on chromosome 17. †proxy for rs6904596 (R^2 = 0.98).

Panel A. Genome-Wide Significant Associations

Subjective Well-Being (SWB, N = 298,420)

SNPID	CHR	BP	EA	EAF	Beta	SE	R^2	p-value	N	Quasi-Repl
rs3756290	5	130,951,750	A	0.24	-0.0177	0.0031	0.011%	9.6×10^{-9}	286,851	
rs2075677	20	47,701,024	A	0.76	0.0175	0.0031	0.011%	1.5×10^{-8}	288,454	DS**
rs4958581	5	152,187,729	T	0.66	0.0153	0.0027	0.011%	2.3×10^{-8}	294,043	DS***

Neuroticism (N = 170,908)

SNPID	CHR	BP	EA	EAF	Beta	SE	R^2	p-value	N	Quasi-Repl
rs2572431#	8	11,105,077	T	0.41	0.0283	0.0035	0.039%	4.2×10^{-16}	170,908	SWB*
rs193236081##	17	44,142,332	T	0.77	-0.0284	0.0045	0.028%	6.3×10^{-11}	151,297	
rs10960103	9	11,699,270	C	0.77	0.0264	0.0038	0.024%	2.1×10^{-10}	165,380	$D^{*}_{23andMe}$
rs4938021	11	113,364,803	T	0.34	0.0233	0.0037	0.024%	4.0×10^{-10}	159,900	$D^{***}_{23andMe}$, SWB*
rs139237746	11	10,253,183	T	0.51	-0.0204	0.0034	0.021%	2.6×10^{-9}	170,908	
rs1557341	18	35,127,427	A	0.34	0.0213	0.0036	0.021%	5.6×10^{-9}	165,579	$D^{**}_{23andMe}$
rs12938775	17	2,574,821	A	0.47	-0.0202	0.0035	0.020%	8.5×10^{-9}	163,283	SWB*
rs12961969	18	35,364,098	A	0.2	0.0250	0.0045	0.020%	2.2×10^{-8}	156,758	
rs35688236	3	34,582,993	A	0.69	0.0213	0.0037	0.019%	2.4×10^{-8}	161,636	
rs2150462	9	23,316,330	C	0.26	-0.0217	0.0038	0.018%	2.7×10^{-8}	170,907	
rs12903563	15	78,033,735	T	0.50	0.0198	0.0036	0.020%	2.9×10^{-8}	157,562	$D^{*}_{23andMe}$,SWB*

Depressive Symptoms (DS, N = 180,866)

SNPID	CHR	BP	EA	EAF	Beta	SE	R^2	p-value	N	Quasi-Repl/Repl
rs7973260	12	118,375,486	A	0.19	0.0306	0.0051	0.029%	1.8×10^{-9}	124,498	$D^{*}_{23andMe}$
rs62100776	18	50,754,633	A	0.56	-0.0252	0.0044	0.031%	8.5×10^{-9}	105,739	$D^{**}_{23andMe}$,SWB*

Panel B. SNPs Identified via Proxy-Phenotype Analyses of SWB Loci with p-value<10^{-4}

Depressive Symptoms in Non-Overlapping Cohorts

SNPID	CHR	BP	EA	EAF	$Beta_{DS}$	SE_{DS}	R^2	p_{DS}	Bonferroni	N_{DS}
rs4346787†	6	27,491,299	A	0.113	-0.023	0.0059	0.011%	9.8×10^{-5}	0.0160	142,265
rs4481363	5	164,483,794	A	0.524	0.014	0.0038	0.009%	3.1×10^{-4}	0.0499	142,265

Neuroticism in Non-Overlapping Cohorts

SNPID	CHR	BP	EA	EAF	$Beta_{neuro}$	SE_{neuro}	R^2	p_{neuro}	Bonferroni	N_{neuro}
rs10838738	11	47,663,049	A	0.49	0.0178	0.0039	0.016%	5.0×10^{-6}	0.0009	131,864
rs10774909	12	117,674,129	C	0.52	-0.0150	0.0039	0.011%	1.2×10^{-4}	0.0203	131,235
rs6904596	6	27,491,299	A	0.09	-0.0264	0.0072	0.012%	2.5×10^{-4}	0.0423	116,335
rs4481363	5	164,474,719	A	0.49	0.0151	0.0040	0.011%	1.9×10^{-4}	0.0316	122,592

[1] Department of Applied Economics, Erasmus School of Economics, Erasmus University Rotterdam, 3062 PA, Rotterdam, The Netherlands
[2] Department of Epidemiology, Erasmus Medical Center, Rotterdam, 3015 GE, The Netherlands
[3] Erasmus University Rotterdam Institute for Behavior and Biology, Rotterdam, 3062 PA, The Netherlands
[4] Department of Biological Psychology, Vrije Universiteit, Amsterdam, 1081 BT, The Netherlands
[5] EMGO+ Institute for Health and Care Research, Amsterdam, 1081 BT, The Netherlands
[6] Saïd Business School, University of Oxford, Oxford, OX1 1HP, UK
[7] Department of Economics, Harvard University, Cambridge, MA 02138, USA
[8] Center for Economic and Social Research, University of Southern California, Los Angeles, CA 90089-3332, USA
[9] Department of Complex Trait Genetics, VU University, Center for Neurogenomics and Cognitive Research, Amsterdam, 1081 HV, The Netherlands
[10] Amsterdam Business School, University of Amsterdam, Amsterdam, 1018 TV, The Netherlands
[11] Psychology, University of Illinois, IL 61820, Champaign, USA
[12] Department of Psychology, University of Minnesota Twin Cities, Minneapolis, MN 55455, USA
[13] COPSAC, Copenhagen Prospective Studies on Asthma in Childhood, Herlev and Gentofte Hospital, University of Copenhagen, Copenhagen, 2820, Denmark
[14] The Novo Nordisk Foundation Center for Basic Metabolic Research, Section of Metabolic Genetics, University of Copenhagen, Faculty of Health and Medical Sciences, Copenhagen, 2100, Denmark
[15] Steno Diabetes Center, Gentofte, 2820, Denmark
[16] Department of Public Health, University of Helsinki, Helsinki, FI-00014, Finland
[17] Centre for Environmental and Preventive Medicine, Wolfson Institute of Preventive Medicine, Queen Mary University of London, London EC1M 6BQ, UK
[18] USDA/ARS Children's Nutrition Research Center, Baylor College of Medicine, Houston, TX 77030, USA
[19] 23andMe, Inc., Mountain View, CA 94041, USA
[20] Centre for Research in Environmental Epidemiology, Institute for Global Health, Barcelona, Spain
[21] Universitat Pompeu Fabra, Barcelona, Spain
[22] CIBER Epidemiología y Salud Pública, Barcelona, Spain
[23] Department of Medical Epidemiology and Biostatistics, Karolinska Institutet, Stockholm, 171 77, Sweden
[24] Institute of Epidemiology II, Helmholtz Zentrum München, German Research Center for Environmental Health, Neuherberg, 85764, Germany
[25] IBS, Unit of Personality, Work and Health, Institute of Behavioural Sciences, P.O. Box 9, 00014 University of Helsinki, Finland
[26] Folkhälsan Research Centre, Helsingfors, FI-00014, Finland
[27] Quantitative Genetics, QIMR Berghofer Medical Research Institute, Brisbane, QLD 4029, Australia
[28] Lifespan Psychology, Max Planck Institute for Human Development, Berlin, 14195, Germany
[29] Max Planck Institute for Molecular Genetics, Department of Vertebrate Genomics, Berlin, 14195, Germany
[30] Estonian Genome Center, University of Tartu, Tartu, 51010, Estonia
[31] Clinical Epidemiology, Leiden University Medical Center, Leiden, 2300 RC, The Netherlands
[32] Public Health and Primary Care, Leiden University Medical Center, Leiden, 2300 RC, The Netherlands
[33] BESC, King Faisal Specialist Hospital and Research Centre, Riyadh 11211, Saudi Arabia
[34] Department of Epidemiology, University of Groningen, University Medical Center Groningen, Groningen, 9700 RB, The Netherlands
[35] Public Health Stream, Hunter Medical Research Institute, New Lambton, NSW 2305, Australia
[36] Faculty of Health and Medicine, University of Newcastle, Newcastle, NSW 2300, Australia
[37] Laboratory of Genetics, National Institute on Aging, Baltimore, MD 21224, USA
[38] Research Centre of Applied and Preventive Cardiovascular Medicine, University of Turku, Turku, 20521, Finland
[39] Department of Clinical Physiology, Turku University Hospital, Turku 20520, Finland
[40] Research Unit of Molecular Epidemiology, Helmholtz Zentrum München, German Research Center for Environmental Health, Neuherberg, 85764, Germany
[41] Department of Psychology, University of Tartu, Tartu 50409, Estonia
[42] Department of Medical Genetics, University of Lausanne, Lausanne, 1005, Switzerland

[43] Swiss Institute of Bioinformatics, Lausanne, 1015, Switzerland
[44] Institute for Medical Informatics, Biometry and Epidemiology, University Hospital of Essen, Essen, 45147, Germany
[45] Icelandic Heart Association, Kopavogur, 201, Iceland
[46] Faculty of Medicine, University of Iceland, Reykjavik, 101, Iceland
[47] MRC Integrative Epidemiology Unit, University of Bristol, Bristol, BS8 2BN, UK
[48] Los Angeles Biomedical Research Institute and Department of Pediatrics, Harbor-UCLA, Torrence, 90505 CA, USA
[49] Institute of Epidemiology and Social Medicine, University of Muenster, Muenster, 48149, Germany
[50] Divisions of Genetics and Rheumatology, Department of Medicine, Brigham and Women's Hospital, Harvard Medical School, Boston, MA 02115, USA
[51] Department of Epidemiology, University of Michigan, Ann Arbor, MI 48104, USA
[52] Department of Gastroenterology and Hepatology, University of Groningen, University Medical Center Groningen, Groningen, 9713 GZ, The Netherlands
[53] Department of Behavioral Sciences, Rush University Medical Center, Chicago, IL 60612, USA
[54] RAND Corporation, Santa Monica, CA 9041-3208, USA
[55] Centre for Cognitive Ageing and Cognitive Epidemiology, University of Edinburgh, Edinburgh, EH8 9JZ, UK
[56] Department of Psychology, University of Edinburgh, Edinburgh, EH8 9JZ, UK
[57] German Socio-Economic Panel Study, DIW Berlin, Berlin, 10117, Germany
[58] Health Economics Research Centre, Nuffield Department of Population Health, University of Oxford, Oxford, OX3 7LF, UK
[59] The Farr Institute of Health Informatics, University College London, London NW1 2DA, UK
[60] Survey Research Center, Institute for Social Research, University of Michigan, Ann Arbor, MI 48104, USA
[61] National Institute on Aging, National Institutes of Health, Baltimore, MD 21224, USA
[62] Institute of Human Genetics, University of Bonn, Bonn, 53127, Germany
[63] Department of Genomics, Life and Brain Center, University of Bonn, Bonn, 53127, Germany
[64] Laboratory of Epidemiology, Demography, National Institute on Aging, National Institutes of Health, Bethesda, MD 20892-9205, United States
[65] Program in Translational NeuroPsychiatric Genomics, Departments of Neurology & Psychiatry, Brigham and Women's Hospital, Boston, MA 02115, USA
[66] Harvard Medical School, Boston, MA 02115, USA
[67] Program in Medical and Population Genetics, Broad Institute of Harvard and Massachusetts Institute of Technology, Cambridge, MA, 02142, USA
[68] Department of Genomics of Common Disease, Imperial College London, London, W12 0NN, UK
[69] Department of Pediatrics, University of Helsinki, Helsinki, Finland
[70] National Institute for Health and Welfare, Helsinki, Finland
[71] Center for Life Course Health Research, University of Oulu; Oulu University Hospital, Oulu, Finland
[72] Institute for Social & Economic Research, University of Essex, Wivenhoe Park CO4 3SQ, UK
[73] Neuroepidemiology Section, National Institute on Aging, National Institutes of Health, Bethesda, MD 20892-9205, USA
[74] LifeLines Cohort Study, University of Groningen, University Medical Center Groningen, Groningen, 9713 BZ, The Netherlands
[75] Department of Internal Medicine, Internal Medicine, Lausanne University Hospital (CHUV), Lausanne, 1011, Switzerland
[76] Molecular Epidemiology, QIMR Berghofer Medical Research Institute, Brisbane, QLD 4029, Australia
[77] Department of Neuroscience, Karolinska Institutet, Retzius Väg 8 171 65 Stockholm, Sweden
[78] Centre for Genomic and Experimental Medicine, Institute of Genetics and Molecular Medicine, University of Edinburgh, Edinburgh, EH4 2XU, UK
[79] Helsinki Collegium for Advanced Studies, University of Helsinki, Helsinki 00014, Finland
[80] Department of Twin Research and Genetic Epidemiology, King's College London, London, SE1 7EH, UK
[81] Centre for Global Health Research, The Usher Institute for Population Health Sciences and Informatics, University of Edinburgh, Edinburgh, EH8 9AG, UK
[82] Information Based Medicine Stream, Hunter Medical Research Institute, New Lambton, NSW 2305, Australia
[83] Department of Neurological Sciences, Rush University Medical Center, Chicago, IL 60612, USA

[84] Lübeck Interdisciplinary platform for Genome Analytics (LIGA), Institutes of Neurogenetics and Integrative & Experimental Genomics, University of Lübeck, Lübeck, 23562, Germany
[85] Neuroepidemiology and Ageing Research Unit, School of Public Health, Faculty of Medicine, Imperial College, London SW7 2AZ, UK
[86] Department of Health Sciences, Community & Occupational Medicine, University of Groningen, University Medical Center Groningen, Groningen, 9713 AV, The Netherlands
[87] Channing Division of Network Medicine, Department of Medicine, Brigham and Women's Hospital, Boston, MA 02115, USA
[88] Istituto di Ricerca Genetica e Biomedica (IRGB), Consiglio Nazionale delle Ricerche, c/o Cittadella Universitaria di Monserrato, Monserrato, Cagliari, 9042, Italy
[89] Department of General Practice and Primary Health Care, University of Helsinki, Helsinki, 00014, Finland
[90] Department of Chronic Disease Prevention, National Institute for Health and Welfare, Helsinki, Finland
[91] Unit of General Practice, University Central Hospital, Helsinki, Finland
[92] Department of Genetics, University Medical Center Groningen, University of Groningen, Groningen. 9700 RB, The Netherlands
[93] Econometric Institute, Erasmus School of Economics, Erasmus University Rotterdam, Rotterdam, 3062 PA, The Netherlands
[94] MRC Human Genetics Unit, Institute of Genetics and Molecular Medicine, University of Edinburgh, Edinburgh, EH4 2XU, UK
[95] Generation Scotland, Centre for Genomics and Experimental Medicine, Institute of Genetics and Molecular Medicine, University of Edinburgh, Edinburgh, EH4 2XU, UK
[96] Department of Psychiatry, Washington University School of Medicine, St. Louis, MO 63110, USA
[97] Centre for Population Health Research, School of Health Sciences and Sansom Institute, University of South Australia, SA5000, Adelaide, Australia
[98] South Australian Health and Medical Research Institute, Adelaide, SA5000, Australia
[99] Population, Policy and Practice, UCL Institute of Child Health, London, WC1N 1EH, UK
[100] Department of Epidemiology and Biostatistics, MRC Health Protection Agency (HPE) Centre for Environment and Health, School of Public Health, Imperial College London, UK
[101] Biocenter Oulu, University of Oulu, Oulu, Finland
[102] Unit of Primary Care, Oulu University Hospital, Oulu, Finland
[103] Institute for Molecular Medicine Finland (FIMM), University of Helsinki, Helsinki, 00014, Finland
[104] Department for Health, THL-National Institute for Health and Welfare, Helsinki, FI-00271, Finland
[105] Department of Epidemiology and Biostatistics, Harvard T.H. Chan School of Public Health, Boston, MA 02115, USA
[106] Department of Social and Behavioral Sciences, Harvard T.H. Chan School of Public Health, Boston, MA 02115, USA
[107] Fimlab Laboratories, Tampere, 33520, Finland
[108] Department of Clinical Chemistry, University of Tampere, School of Medicine, Tampere, 33014, Finland
[109] Genetic Epidemiology, QIMR Berghofer Medical Research Institute, Brisbane, QLD 4029, Australia
[110] Institute of Molecular and Cell Biology, University of Tartu, Tartu, 51010, Estonia
[111] Department of Sociology, University of Oxford, Oxford OX1 3UQ, UK
[112] Department of Psychiatry, University of Groningen, University Medical Center Groningen, 9700 RB, Groningen, The Netherlands
[113] Laboratory of Experimental Cardiology, University Medical Center Utrecht, Utrecht, 3584 CX, The Netherlands
[114] Laboratory of Clinical Chemistry and Hematology, Division Laboratories and Pharmacy, University Medical Center Utrecht, Utrecht, 3584 CX, The Netherlands
[115] Social, Genetic and Developmental Psychiatry Centre, King's College London, De Crespigny Park SE5 8AF, UK
[116] Department of Public Health, Faculty of Medicine, University of Split, Croatia, Split 21000, Croatia
[117] Department of Public Health Sciences, University of Virginia, Charlottesville, VA 22904, USA
[118] Research Unit for Genetic Epidemiology, Institute of Molecular Biology and Biochemistry, Center of Molecular Medicine, General Hospital and Medical University, Graz, Graz, 8010, Austria
[119] Department of Neurology, General Hospital and Medical University Graz, Graz, 8036, Austria
[120] Department of Economics, Oulu Business School, Oulu, Finland

[121] Institute of Preventive Medicine, Bispebjerg and Frederiksberg Hospitals, The Capital Region, Frederiksberg, 2000, Denmark
[122] Department of Epidemiology & Public Health, University College London, London WC1E 7HB, UK
[123] Department of Geriatrics, Florida State University College of Medicine, Tallahassee, FL 32306, USA
[124] Montpellier Business School, Montpellier, 34080, France
[125] Panteia, Zoetermeer, 2715 CA, The Netherlands
[126] Department of Psychiatry, Erasmus Medical Center, Rotterdam, 3015 GE, The Netherlands
[127] Department of Child and Adolescent Psychiatry, Erasmus Medical Center, Rotterdam, 3015 GE, The Netherlands
[128] Department of Internal Medicine, Erasmus Medical Center, Rotterdam, 3015 GE, The Netherlands
[129] Queensland Brain Institute, The University of Queensland, Brisbane, QLD 4072, Australia
[130] The University of Queensland Diamantina Institute, The Translational Research Institute, Brisbane, QLD 4102, Australia
[131] Department of Sociology, New York University, New York, NY 10012, USA
[132] School of Medicine, New York University, NY 10016, New York, USA
[133] Department of Economics, Stockholm School of Economics, Stockholm, 113 83, Sweden
[134] Bioethics Program, Union Graduate College - Icahn School of Medicine at Mount Sinai, Schenectady, NY 12308, USA
[135] New York Genome Center, New York, NY 10013, USA
[136] Department of Biological Sciences, Columbia University, 600 Fairchild Center, New York, NY 10027, USA
[137] Center for Economic and Social Research, University of Southern California, Los Angeles, CA 90089-3332, USA
[138] Neuroscience Campus Amsterdam, Amsterdam, The Netherlands
[139] Department of Economics, New York University, New York, NY 10012, USA
[140] Research Institute for Industrial Economics, Stockholm, 10215, Sweden